On the Road to the Higgs: Evidence for Semi-leptonic WW / WZ Decays at DØ

Wade Fisher

Fermilab

On behalf of the DØ Collaboration

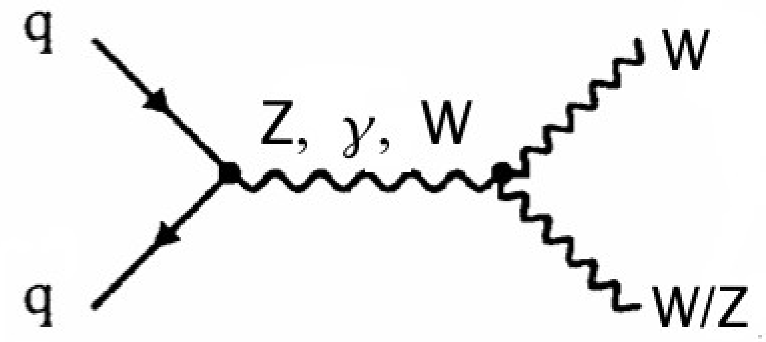
Outline

- ✗ Motivation for diboson studies in WW/WZ→lνjj
- ✗ Data sample and event selection
- ✗ Studies of systematic uncertainties
- ✗ Cross section measurement and statistical significance
- ✗ Multivariate event classification

Dibosons at the Tevatron

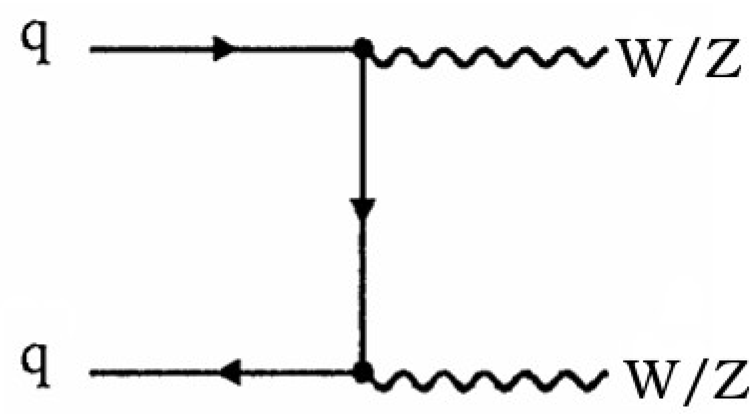
- Studies of electroweak (EW) vector boson production are an important aspect of the Tevatron physics program.
  - The EW group structure is central to the Standard Model: $\mathbf{SU(2)_L \otimes U(1)_Y}$

s - channel



- Potential for new physics is manifest in:
  - Precision measurements of mass and EW parameters
  - Relationships between the masses of the W and Z
  - Increased cross sections or changes in kinematics

t - channel



Tevatron Measurements

- Diboson measurements at the Tevatron are progressing quickly and new final states are frequently appearing
  - ⇒ Currently all diboson measurements are in fully leptonic final states

- Right: total cross sections of Tevatron EW results

σ(WW):

DØ: $13.8^{+4.6}_{-4.0}(stat+syst)\,pb$

CDF: $13.6 \pm 3.1\,(stat+syst)\,pb$

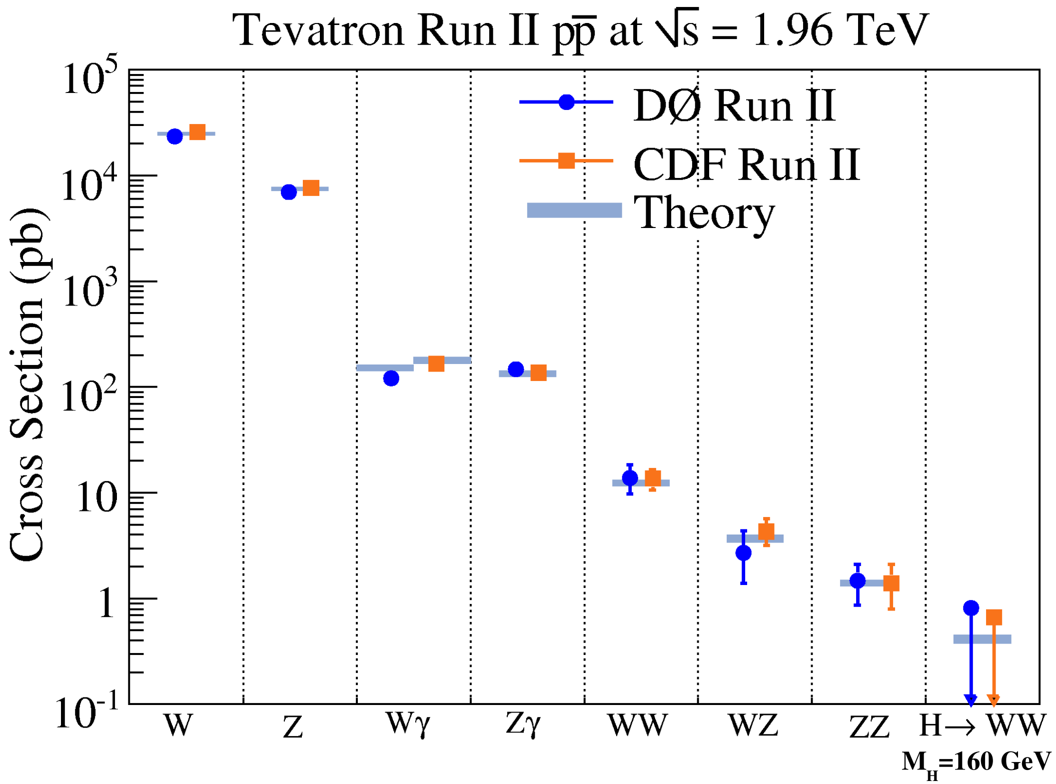
Theory: $12.4 \pm 0.8\,pb$

σ(WZ):

DØ: $2.7^{+1.7}_{-1.3}(stat+syst)\,pb$

CDF: $4.3^{+1.4}_{-1.1}(stat+syst)\,pb$

Theory: $3.7 \pm 0.3\,pb$



ZZ Observation by DØ

- Just two months ago, DØ announced observation of ZZ production in $ZZ \to ll\nu\nu$ and $ZZ \to llll$ (*4 charged leptons*) final states

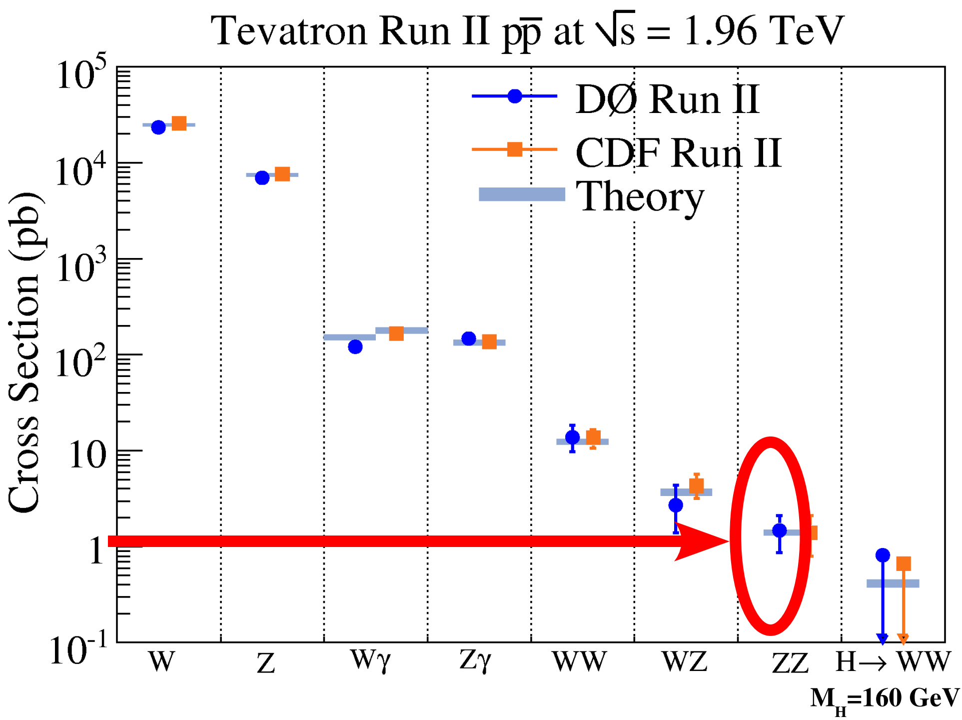
- Cross section measured to be

$\sigma(ZZ) = 1.60 \pm 0.63$ (stat) ± 0.16 (syst) pb

$\sigma(SM) = 1.4 \pm 0.1$ pb

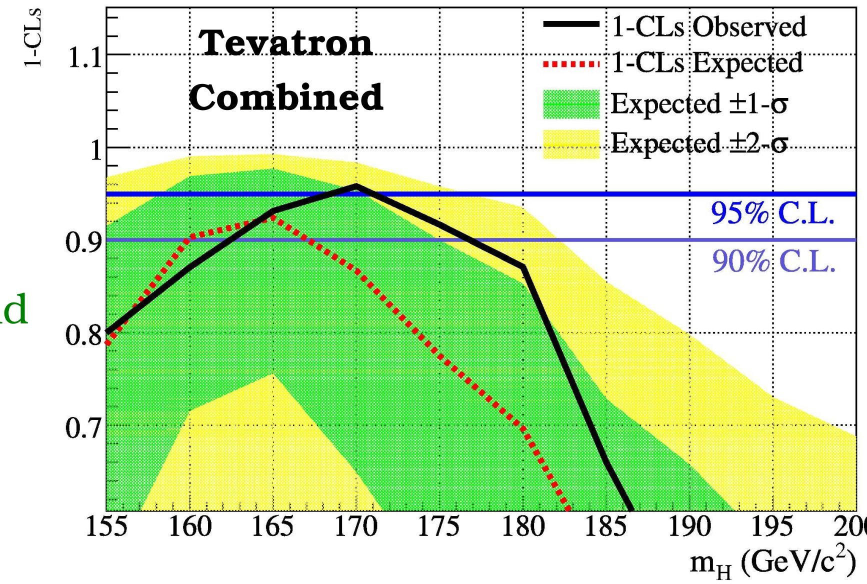
[J.M. Campbell and R.K. Ellis PRD 60, 113006 (1999)]

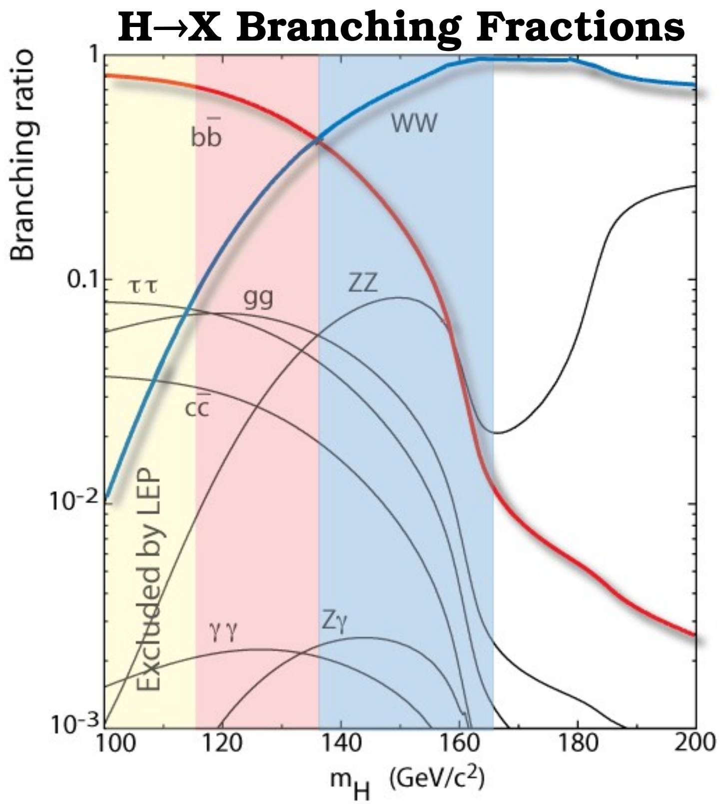
- Significance of the measurement was calculated to be **5.7 standard deviations** above background

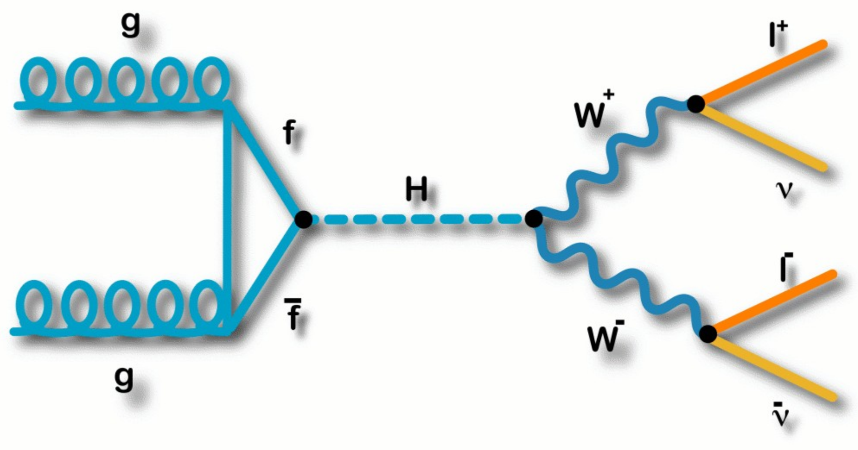


Connecting with Higgs Search

- H→WW is the dominant decay mode for high mass Higgs bosons (M_H>135 GeV)
  - Drives current exclusion limits
  - Direct WW production is dominant background
  - Essential to understand this background

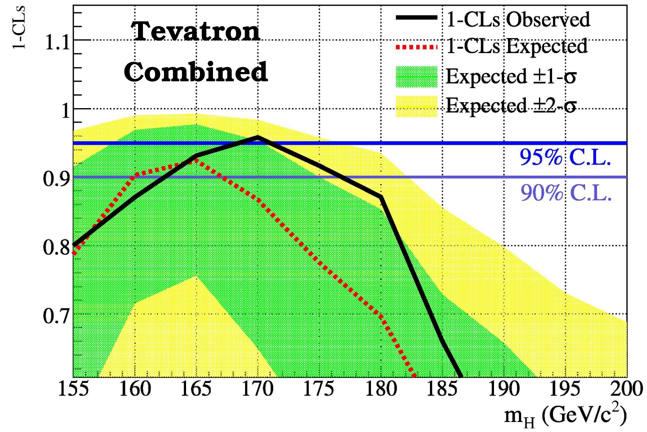




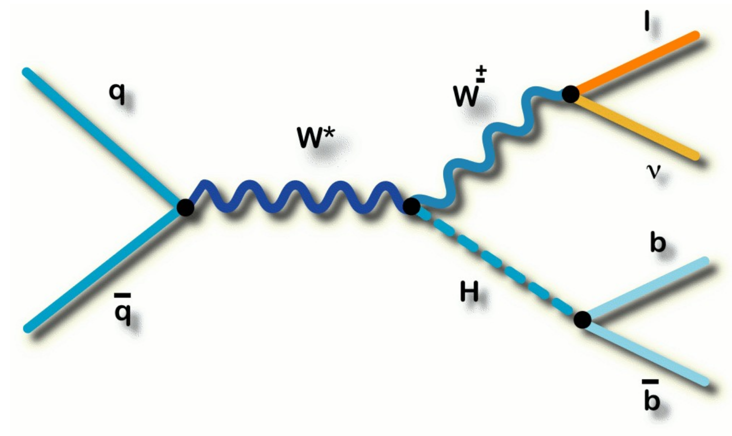


Connecting with Higgs Search

- H→WW is the dominant decay mode for high mass Higgs bosons (M_H>135 GeV)
  - Drives current exclusion limits
  - Direct WW production is dominant background
  - Essential to understand this background



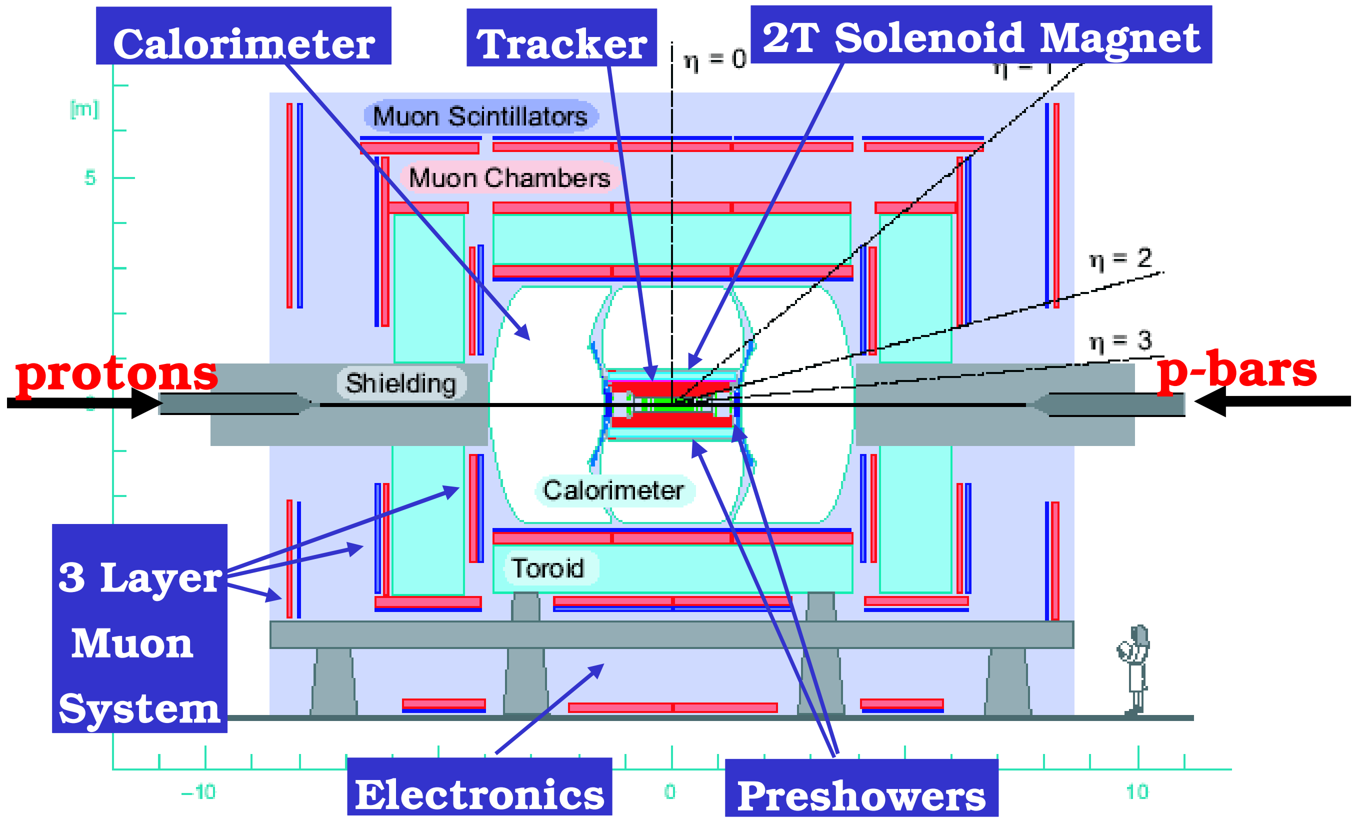
- WH→lνbb is a promising search channel for a low mass Higgs (M_H<135 GeV)
  - Similar final state to WW/WZ→lνjj
  - Similar challenges
    - Small signal in a large background
    - **S/B ⇒ WH: 1.2% WW/WZ: 2.9%**
  - Can test similar analysis techniques
    - Multivariate classifiers, statistical treatment



WW/WZ→lνjj represents a valuable proving ground for analysis techniques used in the Tevatron Higgs search


The
DØ
Collaboration
18 Countries
90 Institutions
554 Scientists

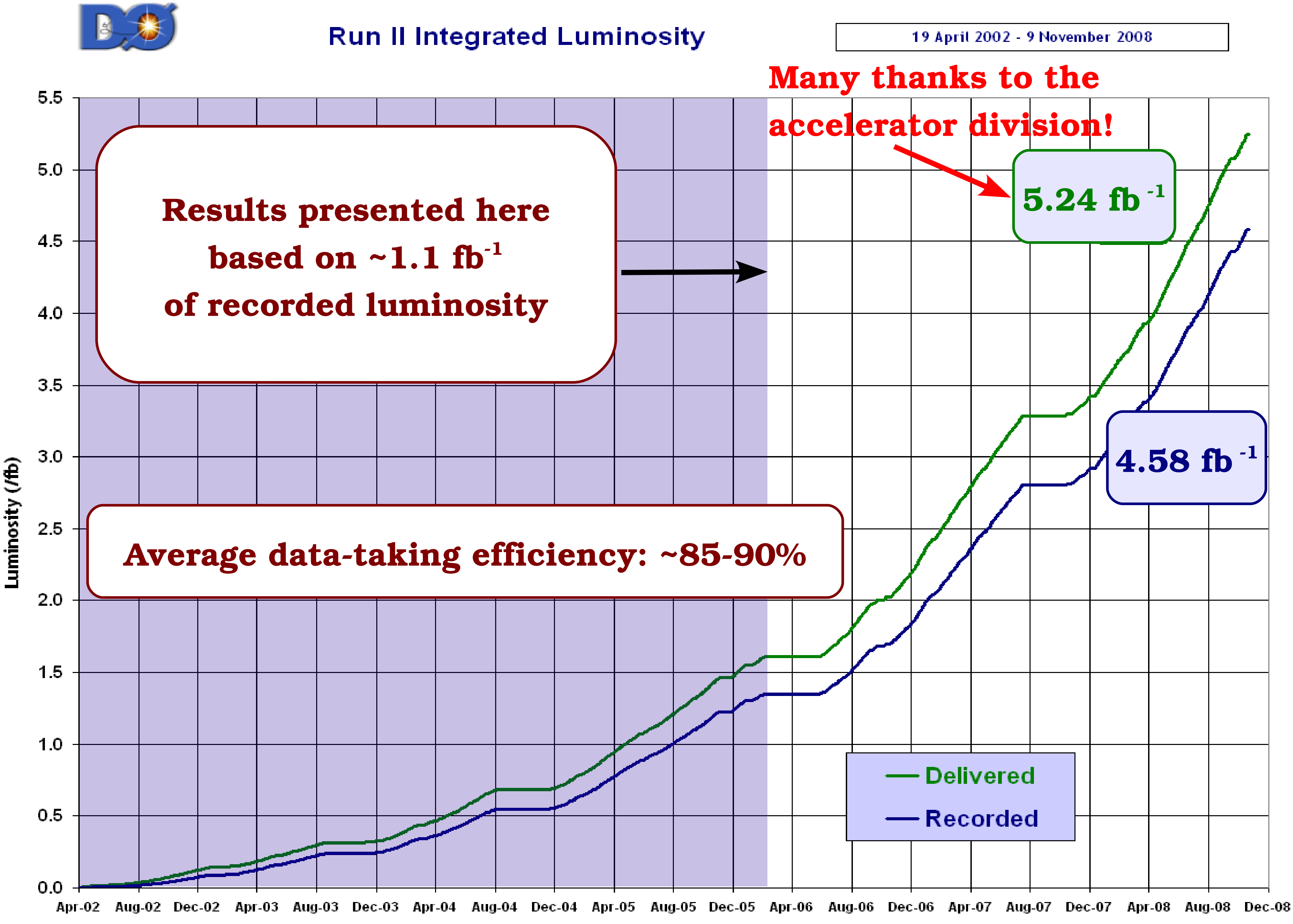
The DØ Detector



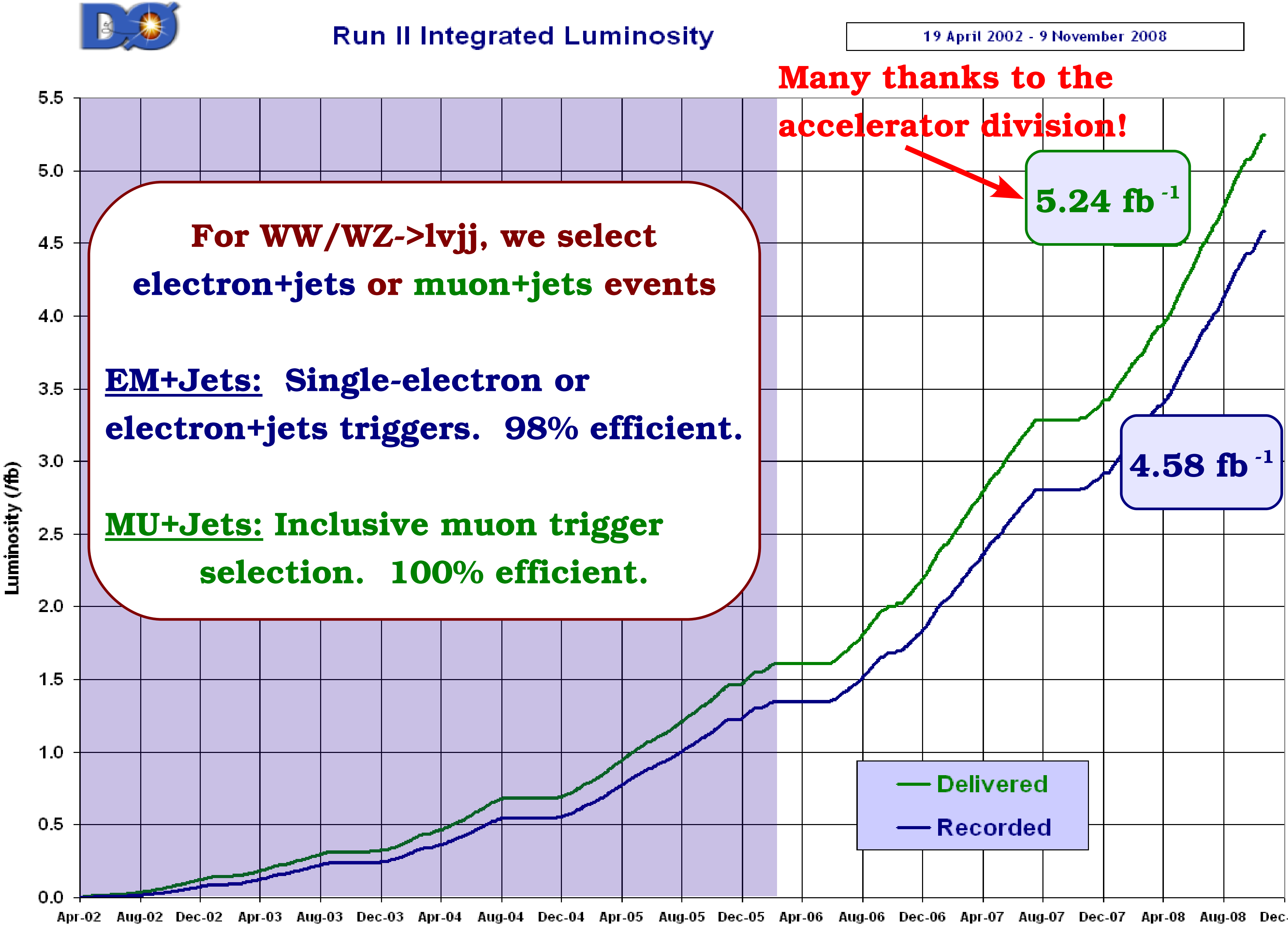
- ✗ Silicon microstrip vertex detector
- ✗ Scintillating fiber tracker
- ✗ Uranium / liquid argon calorimeter
- ✗ Wire chamber + scintillation counter muon detector system
- ✗ 2T solenoid magnet & 1.8T toroid magnet

Angular Coverage	$\lvert\eta\rvert$
Muon ID	~2
Tracking	~2.5
EM / Jet ID	~4

Dataset


Run II Integrated Luminosity
19 April 2002 - 9 November 2008
Many thanks to the accelerator division!
5.24 fb -1
4.58 fb -1
Results presented here based on ~1.1 fb-1 of recorded luminosity
Average data-taking efficiency: ~85-90%
Luminosity (/fb)
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Apr-02
Aug-02
Dec-02
Apr-03
Aug-03
Dec-03
Apr-04
Aug-04
Dec-04
Apr-05
Aug-05
Dec-05
Apr-06
Aug-06
Dec-06
Apr-07
Aug-07
Dec-07
Apr-08
Aug-08
Dec-08
Delivered
Recorded

Dataset



For WW/WZ->lvjj, we select electron+jets or muon+jets events

EM+Jets: **Single-electron or electron+jets triggers. 98% efficient.**

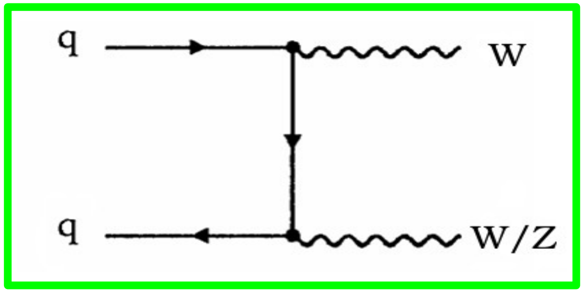
MU+Jets: **Inclusive muon trigger selection. 100% efficient.**

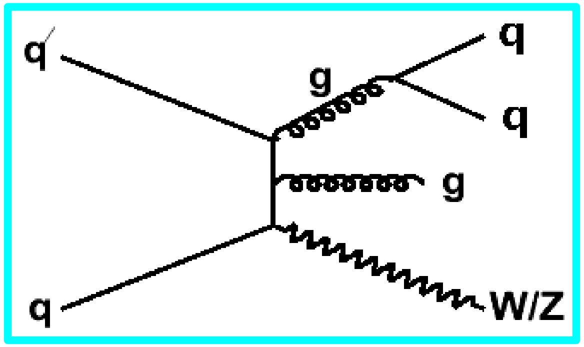
Simulated Samples

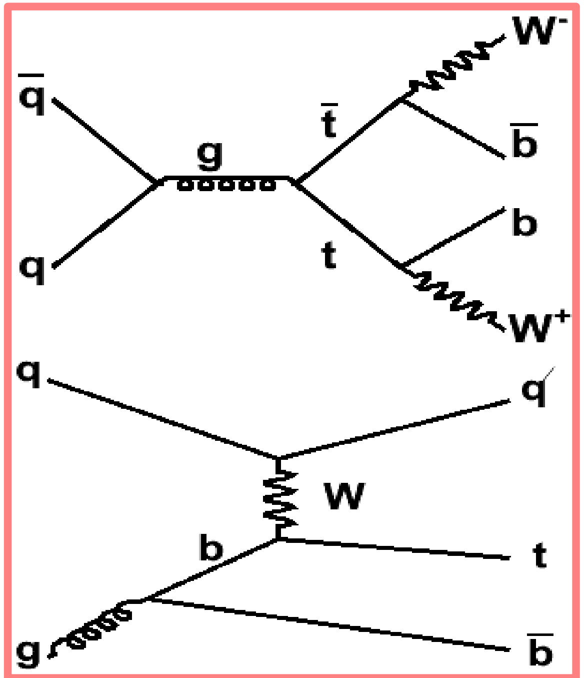
- Nearly all event sources are generated via Monte Carlo with a full simulation of detector response
  - W/Z to tau decays and tau cascade decays are included

Event Source	Generator	σ(SM) / σ(WW) = 12.4 pb
WW	Pythia	1.0
WZ	Pythia	0.3
ZZ	Pythia	0.1
W+light flavor jets	Alpgen	800
W+heavy flavor jets	Alpgen	30
Z+light flavor jets	Alpgen	30
Z+heavy flavor jets	Alpgen	1
Double-Top	Alpgen	0.6
Single-Top	Comphep	0.2

- The rate and distributions of multijet events, in which jets are misidentified as leptons, are determined from data
  - The multijet contributions are corrected for misidentified events in simulated samples.

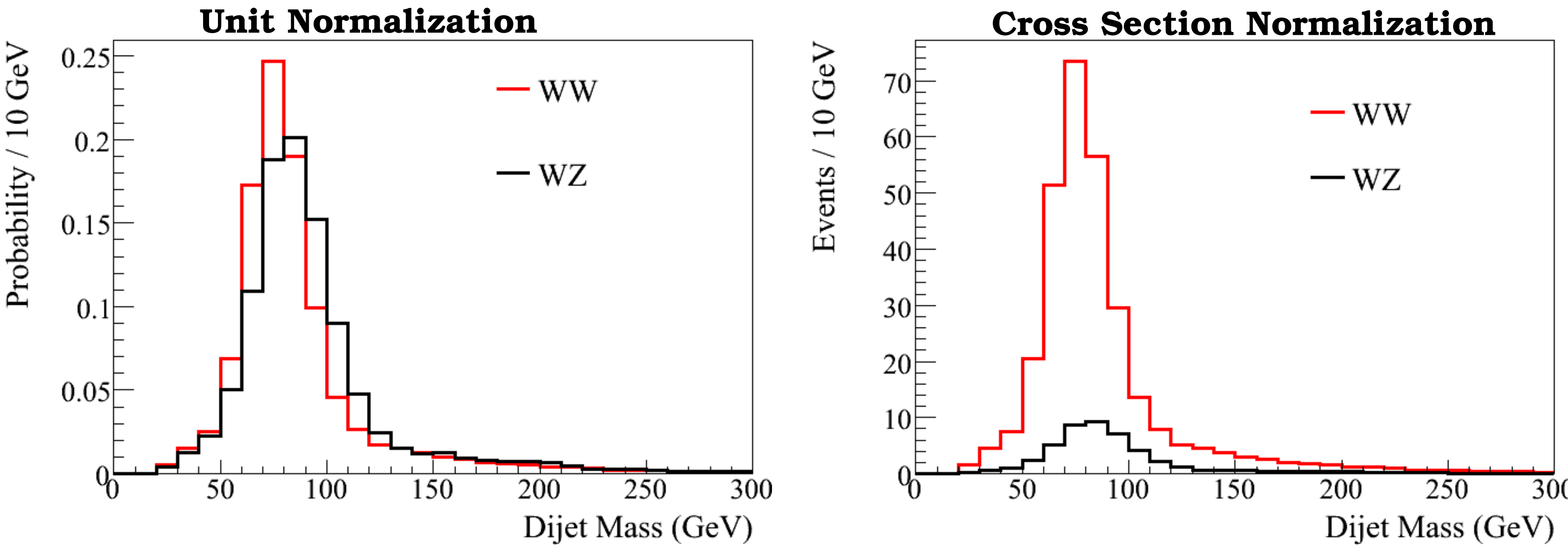






WW vs WZ

- We treat events from WW and WZ as indistinguishable signals
  - Largely due to insufficient dijet mass resolution: ~10 GeV difference in mass
  - Cascade decays of heavy quarks in Z→bb contain neutrinos, thus reducing reconstructed dijet mass in these events. **Final mass difference: ~7 GeV**
- Consider the relative selection efficiency for WW vs WZ
  - WW(WZ)→lνjj branching fraction: ~28.5 (14.2)%
  - WW(WZ)→lνjj $\sigma\times$BR: ~3.5 (0.5) pb



Lepton Selection

✗ Selecting **electrons** and **muons**

p_T > 20 GeV **$|\eta_{EM}|$<1.1** **$|\eta_{MU}|$<2.0**

Spatial match to a central track

Veto events with multiple leptons

Electrons:

Require calorimeter showers consistent with electromagnetic interactions

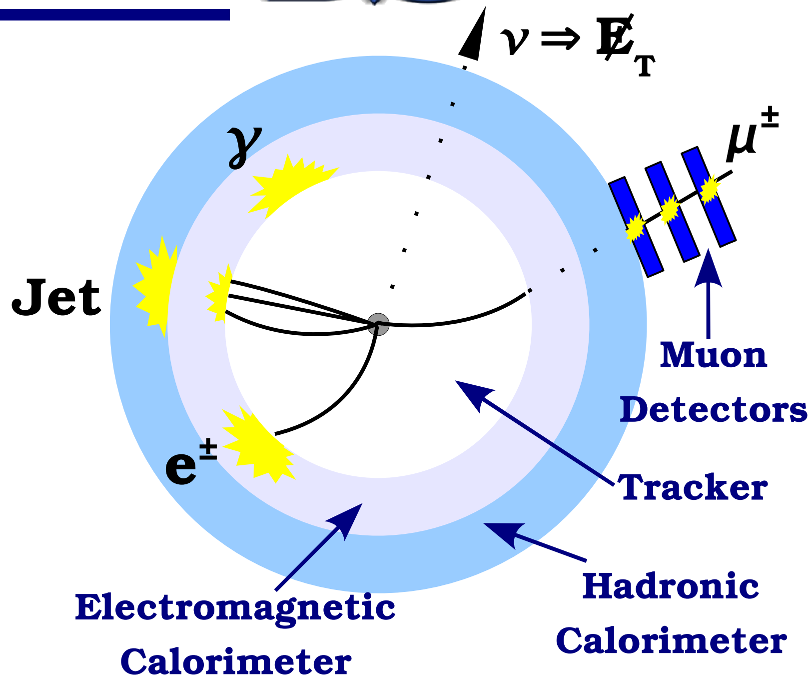
Calorimeter energy is clustered in radial cones of $\Delta R = ((\Delta\phi)^2 + (\Delta\eta^2))^{1/2} < 0.4$

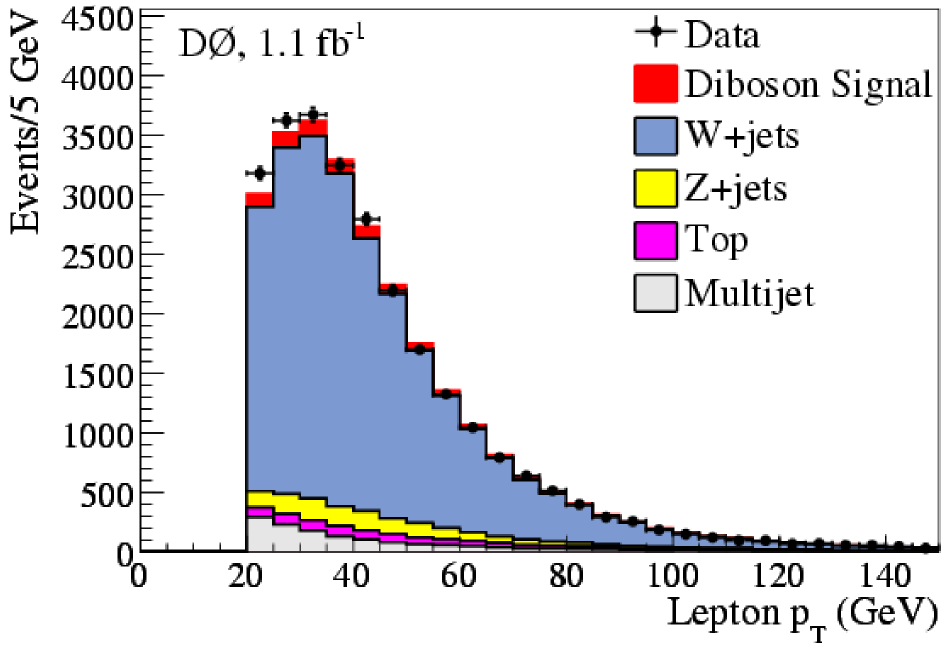
Require that 90% of energy is deposited in the EM calorimeter

Muons:

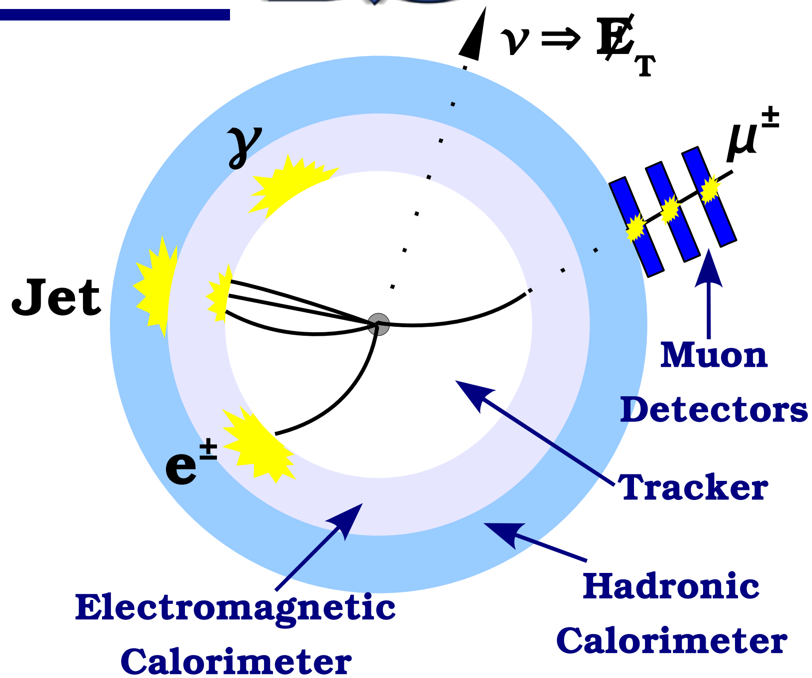
Must have hits in at least three muon detector layers

Signature must be isolated in both the tracker and the calorimeter



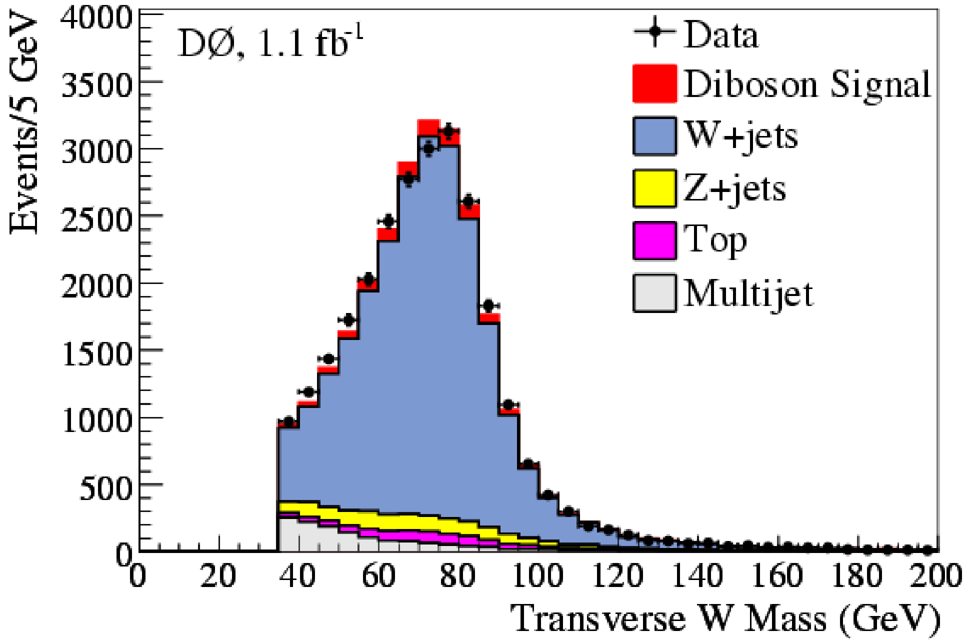


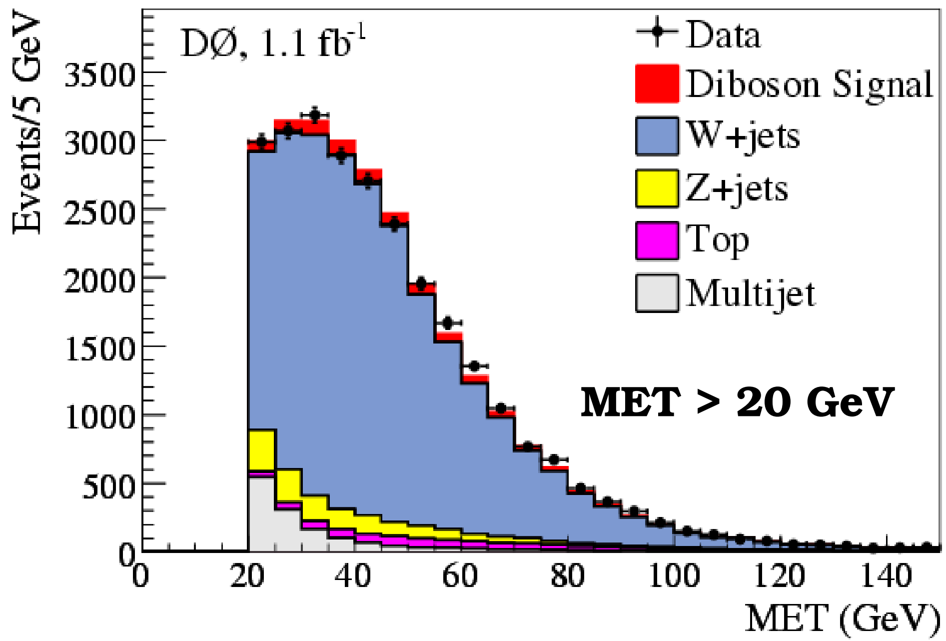
$W \to l\nu$ *Selection*



- Once we have a good lepton, we want events consistent with $W \to l\nu$ decay
  - Undetected neutrino manifests as an imbalance in transverse momentum: "missing" transverse energy (MET)
  - To reduce multijet backgrounds, we require **$M_T^W > 35$ GeV**.

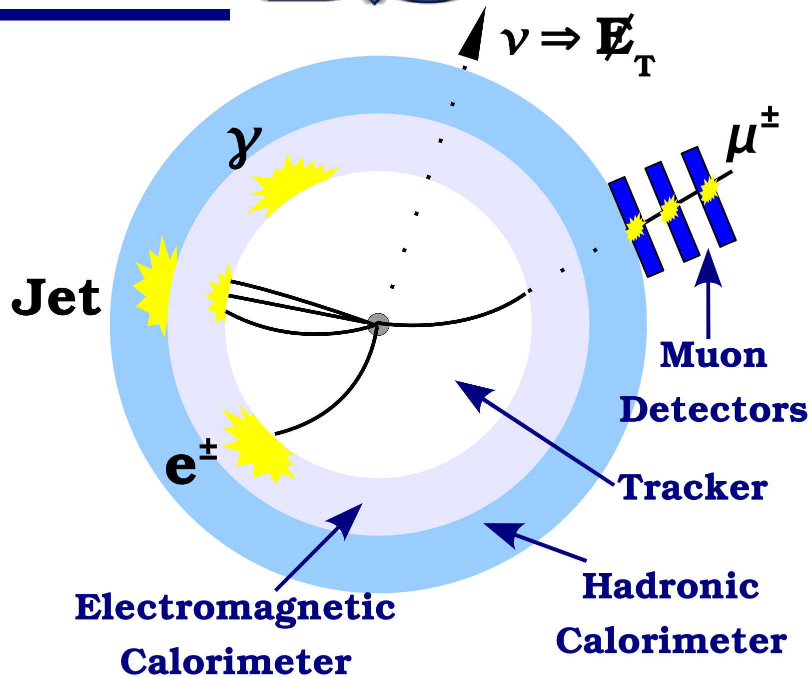
$$M_T^W = \sqrt{2\ p_T(l^{\pm})\ p_T(\nu)\ (1-\cos(\phi(l^{\pm})-\phi(\nu)))}$$

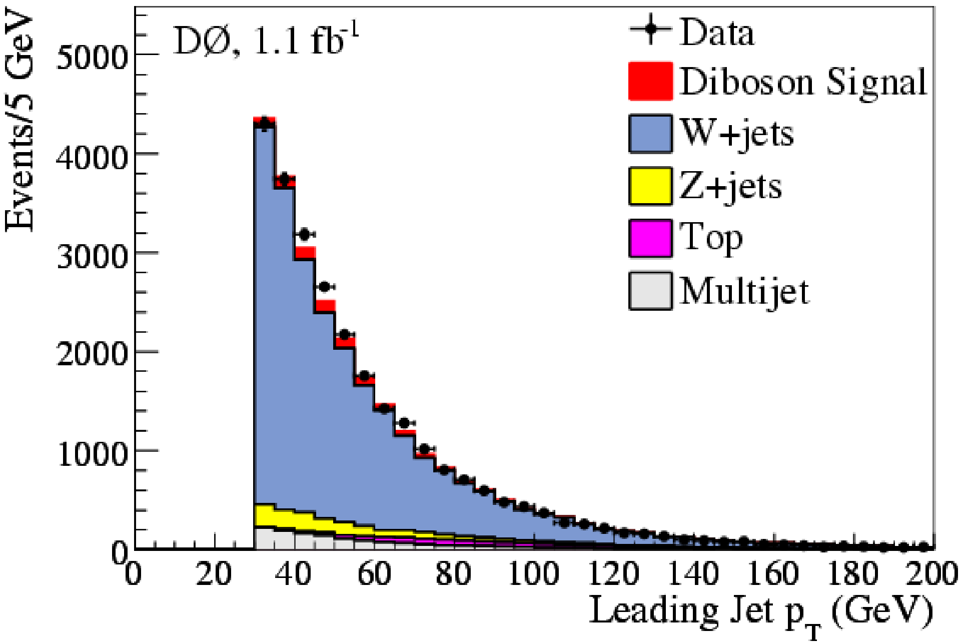


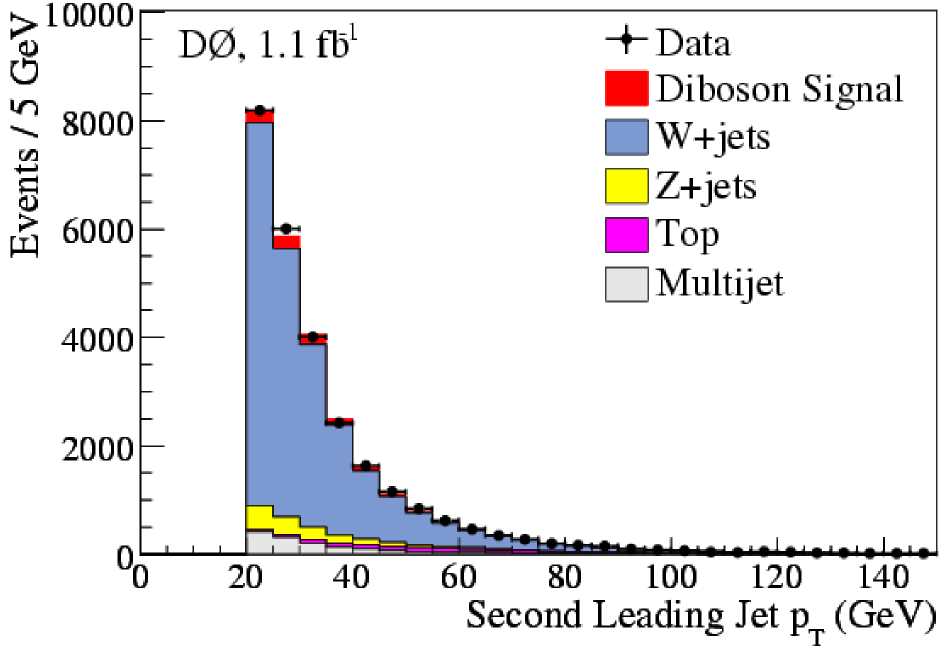


W / Z→qq Selection

- Quark jets arising from W/Z→qq decays are high-energy and relatively central
  - Cluster energy in cones of **ΔR < 0.5**
  - Calorimeter signature must be inconsistent with electron signatures
  - **$|\eta_{JET}|$<2.5** *No veto on # jets*
  - Highest jet **p_T > 30 GeV**
  - 2nd leading jet **p_T > 20 GeV**







Monte Carlo Corrections

- ✗ The event selection process includes efficiency and kinematic corrections for known data/Monte Carlo differences

  Z pT: The transverse momentum of Z bosons in Z+jets events is corrected to measurements in data.

  Multiple Parton Interactions: The simulation of multiple parton interactions in beam collisions is tuned to data.

  Lepton and Jet Identification: Percent-level corrections. Often arise from changes in real detector efficiency during running period.

  Trigger selection: Trigger efficiencies are measured in data and propagated to simulated samples.

  Luminosity profile: The instantaneous luminosity profile of the simulation is matched to data. Helps to properly model minimum bias effects.

  Beam z-position profile: The longitudinal profile of the beam interaction region is matched to data. Impacts angular and energy calculations.

Event Yields

- ✗ Following selection and standard corrections, we may evaluate the expected and observed number of events in each selection
- ✗ The dominant background (~85%) is W+jets. Signal / Background ~ 3%.

Two consequences:

We cannot rely upon the predicted W+jets cross section. **Must determine in data.**

We must study all systematic effects on the background **of the order ~1-2%.**

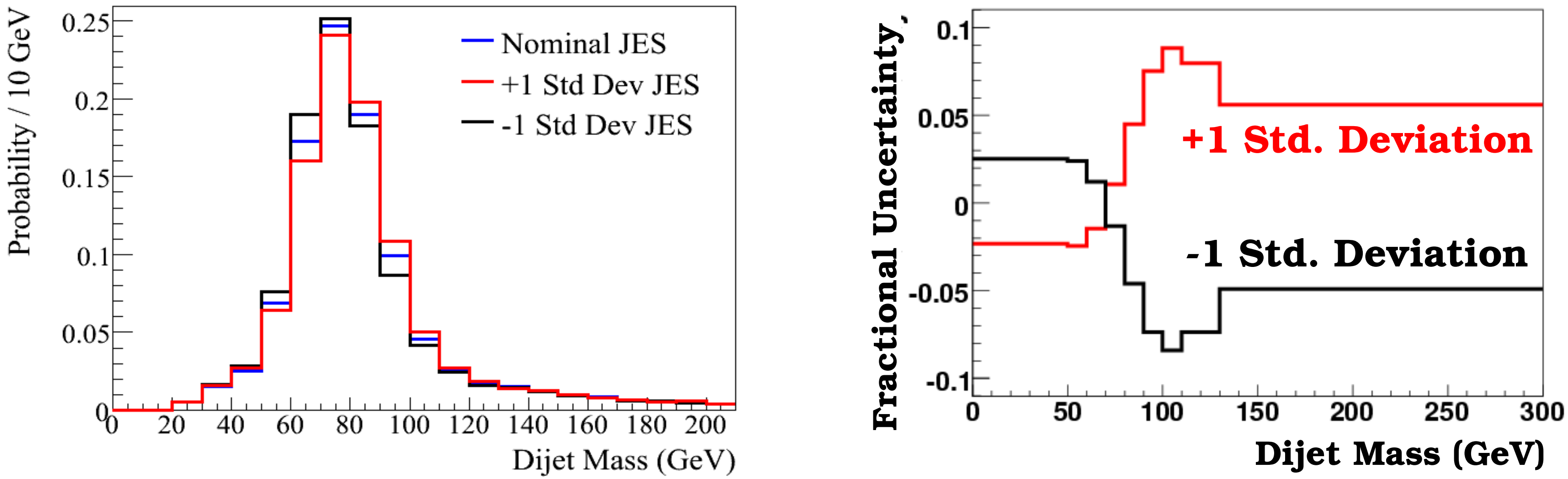
Expected and Observed Numbers of Events

	$e\nu jj$ channel	$\mu\nu jj$ channel
Luminosity	1067 pb^{-1}	1074 pb^{-1}
WV	357.5 ± 2.3	415.8 ± 2.7
W+light flavor jets	8158 ± 72	9681 ± 84
W+heavy flavor jets	2060 ± 26	2319 ± 28
Z+jets	406 ± 13	1237 ± 20
$t\bar{t}$ + single top	463.3 ± 2.2	438.0 ± 2.2
Multijet	825 ± 11	327.0 ± 9.6
ZZ	2.99 ± 0.14	11.53 ± 0.28
Total predicted	12272 ± 78	14428 ± 91
Data	12473	14392

Systematic Uncertainties

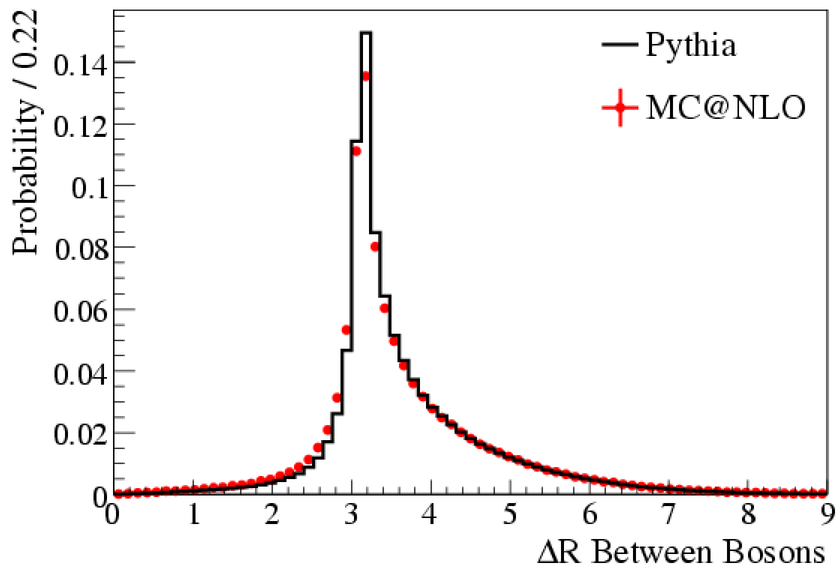
- With an expected signal/background of ~3%, we need to study sources of systematic uncertainty on the level of 1-2%.
- We consider sources of systematic uncertainty at every level of our selection:
  - Luminosity, Trigger efficiency, lepton selection efficiency
  - Jet identification, energy scale, and energy resolution
  - Background cross sections, background modelling effects
- We consider sources of systematic uncertainty that impact both kinematic variable shape and those that have a flat response (eg, Luminosity)
  - Systematic effects are rarely symmetric ⇒ require careful measurement of each source

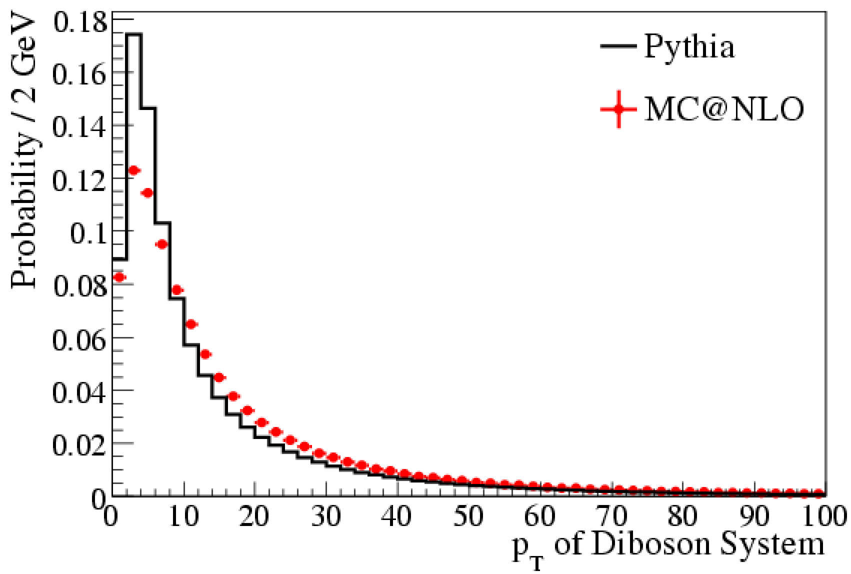
Example: Jet energy scale uncertainty for signal events

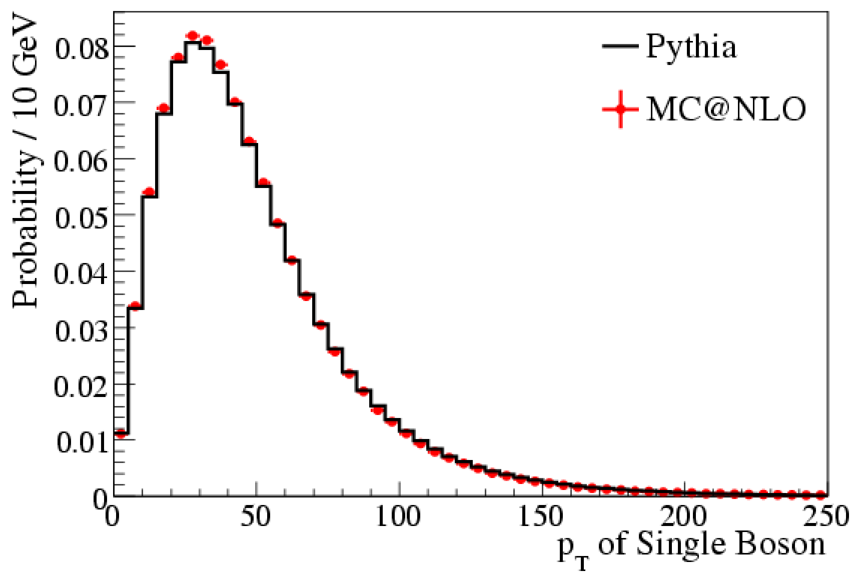


Signal Modeling

- We simulate diboson signal events at leading order (LO) using Pythia
  - Comparisons with NLO using MC@NLO indicate differences in several acceptance-related variables
  - To compensate, we correct the Pythia diboson system at generator level to match MC@NLO
  - Correlations amongst 1-D variables are handled via a 3-D correction model



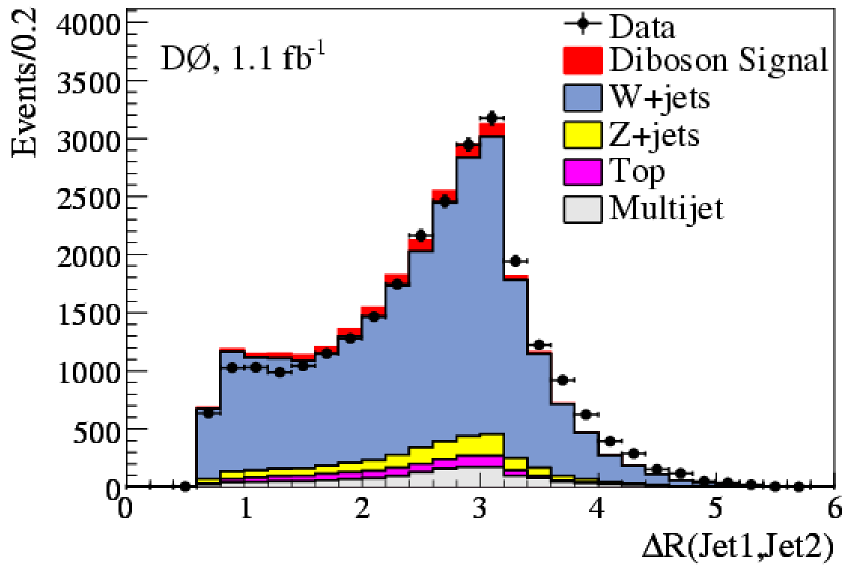


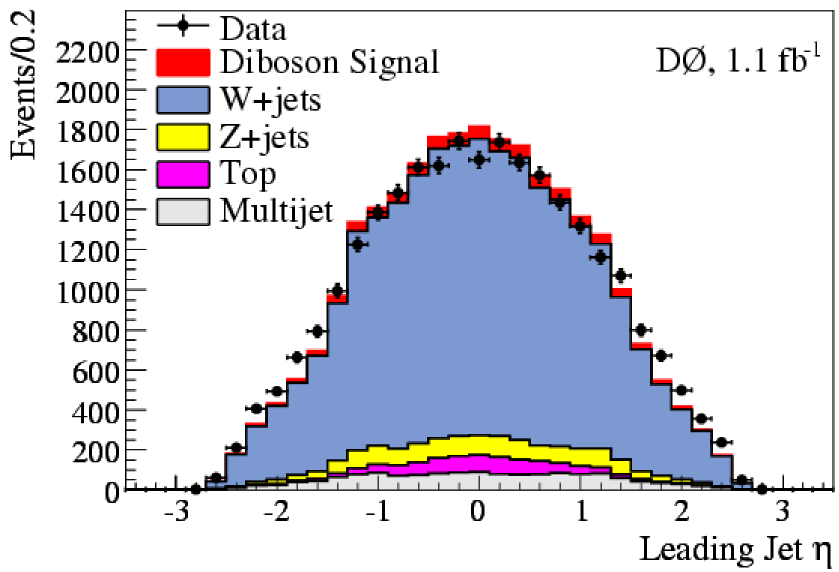


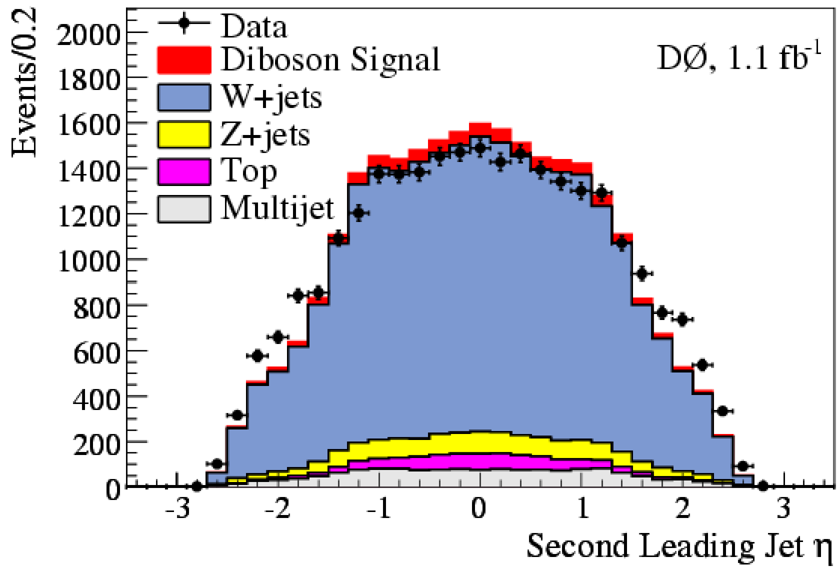
MC@NLO: http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO/

Angular Distributions

- Following event selection, we observe differences between data and MC in the **angular distributions of jets**
  - Differences of this magnitude must come from the dominant V+jets background
  - Recall, jet pT distributions well-modeled
- Detailed studies conclude that these effects arise due to the relative angles of low p_T jets
  - Alpgen generator tuning, Pythia tune, Pythia/Alpgen interface





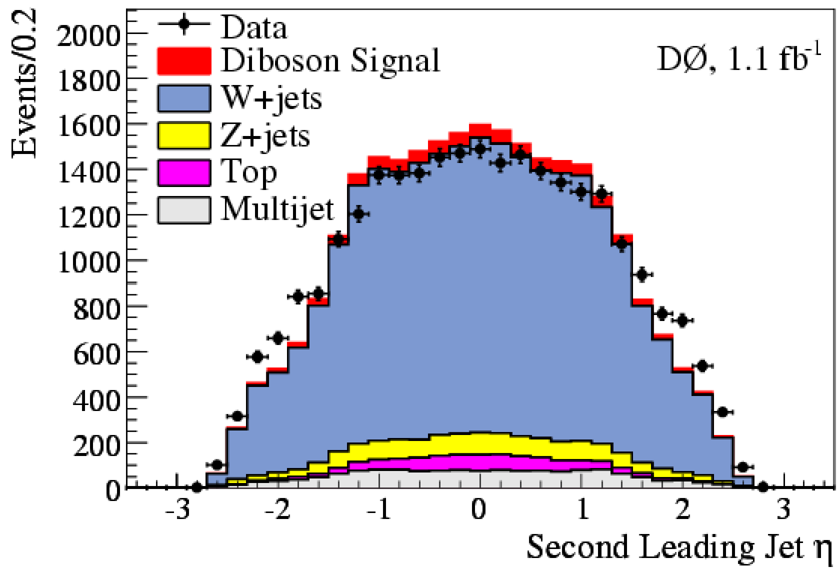


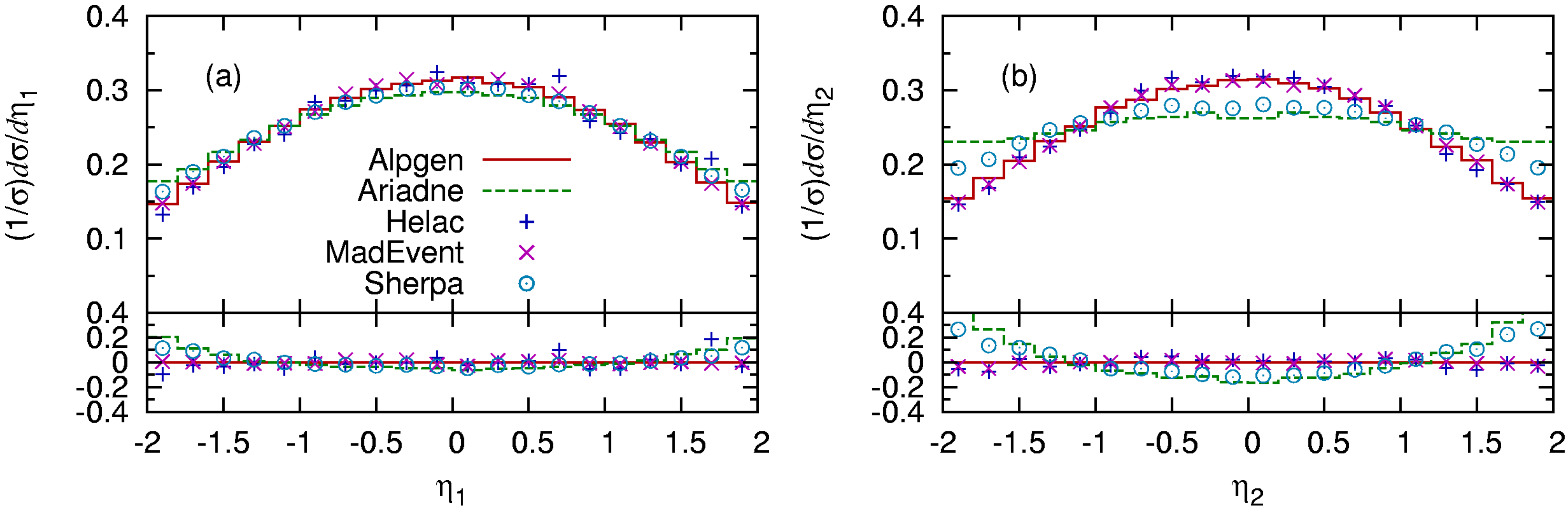
Angular Distributions II

- Studies of these effects in a range of generators reveal similar modeling effects in jet angular distributions

  Size and distributions of the effects nearly identical to that found in the data

arXiv/hep-ph:0706.2569





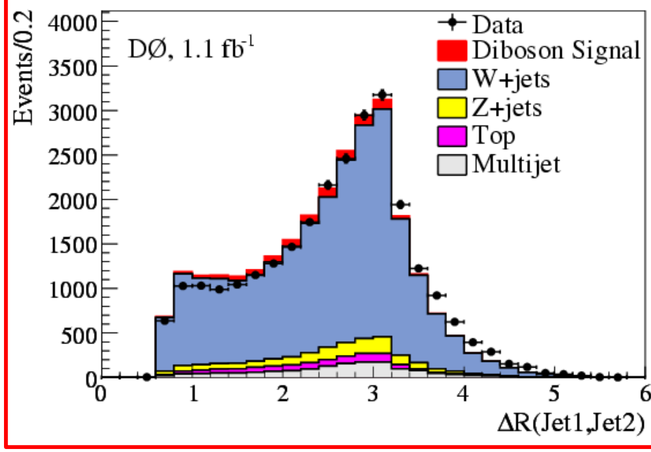
Angular Distributions III

✗ We correct at the event-level using correction functions derived from the relative data/MC shapes

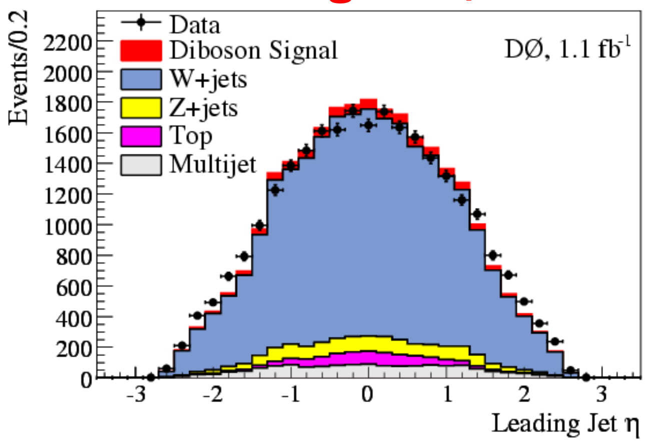
Subtract non-W/Z+jets backgrounds from data and compare to W/Z+jets MC

Before Corrections

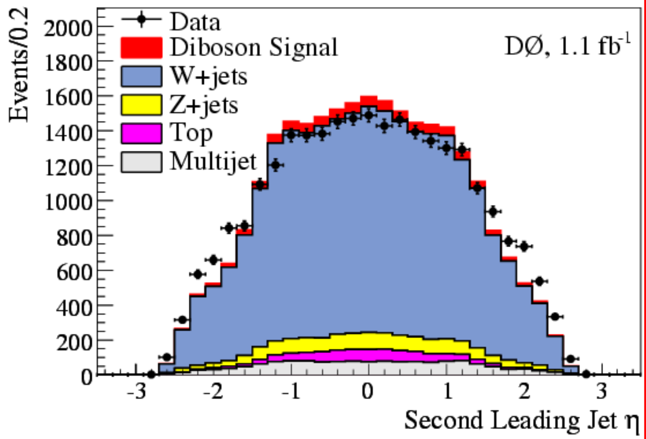
ΔR(jet1, jet2)



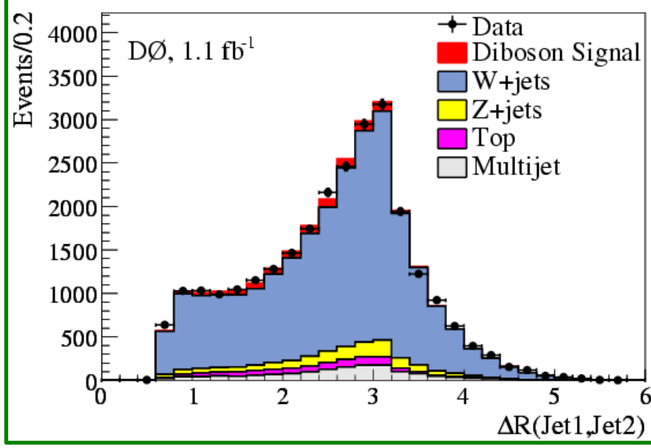
Leading Jet η

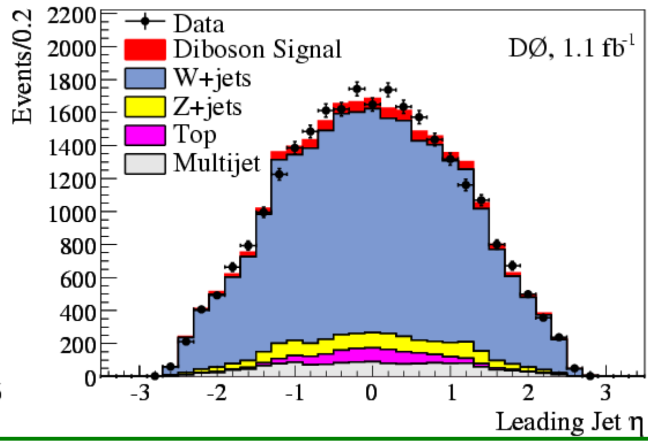


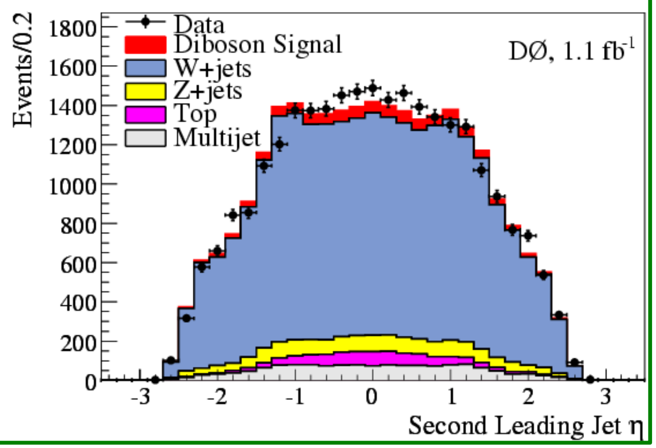
2nd Leading Jet η



After Corrections







Alpgen Modeling

- We've now observed that the event modeling via Alpgen needs scrutiny
  - To ensure proper systematic uncertainty coverage, we next moved to look at uncertainties related to **Alpgen modeling parameters**

Renormalization scale: Dynamic renormalization scale of

$\mathbf{Q^2 = M_W^{\ 2} + \Sigma\ (jet\ p_T)^2}$ can be varied by a constant factor.

kT Scale Factor: Alpgen's scale factor for α_s at each decay vertex.

Parton-matching cluster pT threshold: The minimum pT for jet clusters that are used for the MLM jet-parton matching procedure.
Alpgen authors recommend using the generator-level jet pT cut + 20% (or 5 GeV if larger)

Parton-matching clustering radius size: The size of the jet cone used when creating jet clusters for the MLM jet-parton matching procedure.
Alpgen authors recommend using the generator-level cut.

Parton-Jet Matching

- ✗ While simulating W/Z + N-jets, we need to get the following correct
  1) The inclusive cross section
  2) The relative cross sections for exclusive N-jet final states

We simulate by pairing Alpgen (LO matrix elements) and Pythia (parton showers)

Matrix Elements

1) Fixed order, parton level
2) Accurate description of the hard process
3) Limited number of partons
4) Needed for N-jet description

Parton Showers

1) Resummation of large logs
2) No limit on the number of partons
3) Needed for a realistic description of the final state in the detector

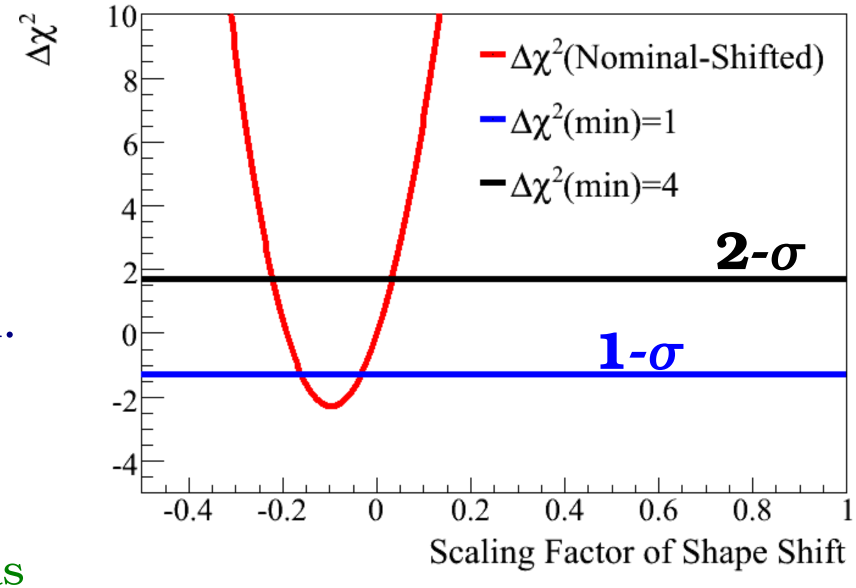
- ✗ These are complementary processes and we need to combine them.

The problem: Double-counting of final states due to jets from showering

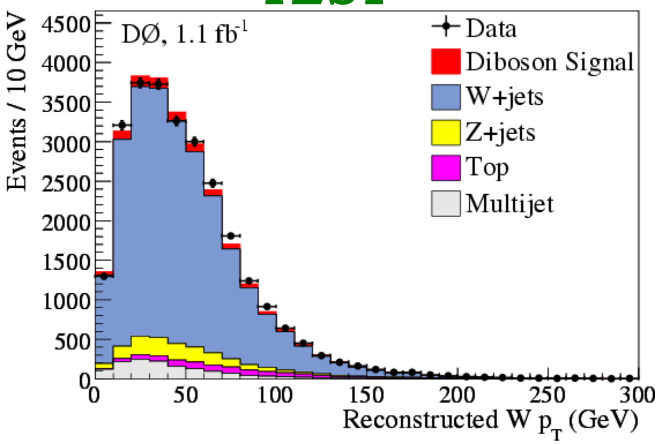
MLM parton-jet matching algorithm (Alpgen): Cluster the showered partons into cone jets. Keep events only if each jet is matched to just one parton.

Alpgen Modeling II

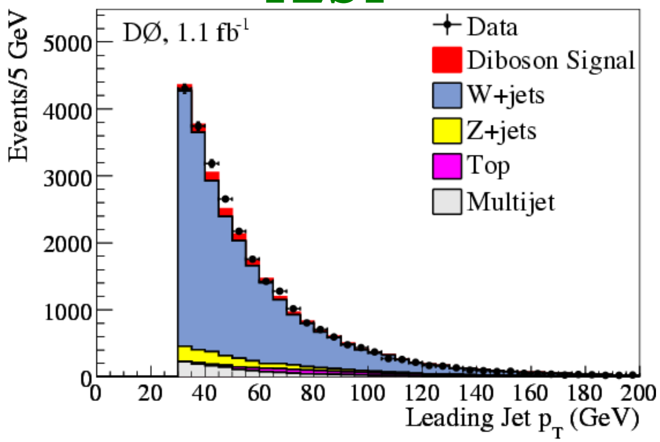
- ✗ Need to test two things: data/MC agreement and systematic uncertainty
  - Prescription: Map $\Delta\chi^2$ as a function of the change in each parameter we test
- ✗ Ultimately we care about the dijet mass, so we don't want to bias that distribution.
  - ⇒ Test **p_T of $W \to l\nu$** and **leading jet p_T**, first removing **$55 < M_{JJ} < 110$ GeV**
  - ⇒ Ignore any change in total normalization, as this will be handled separately



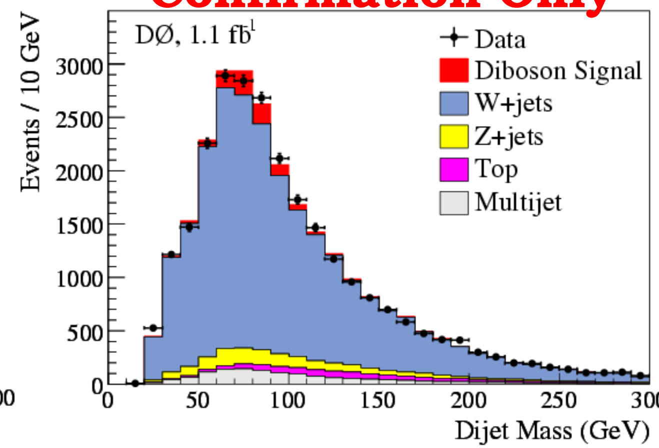
TEST



TEST

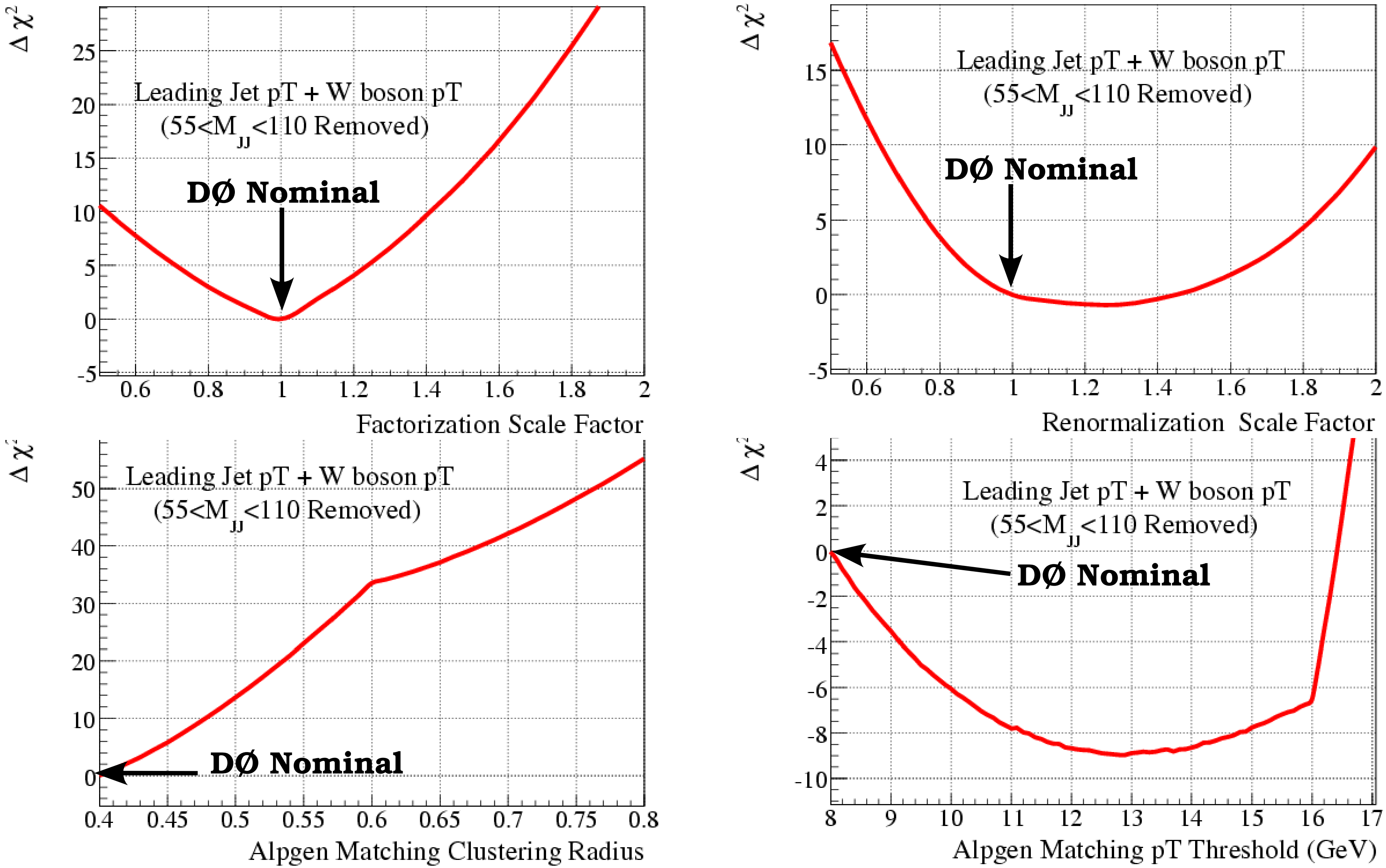


Confirmation Only



Alpgen Modeling III

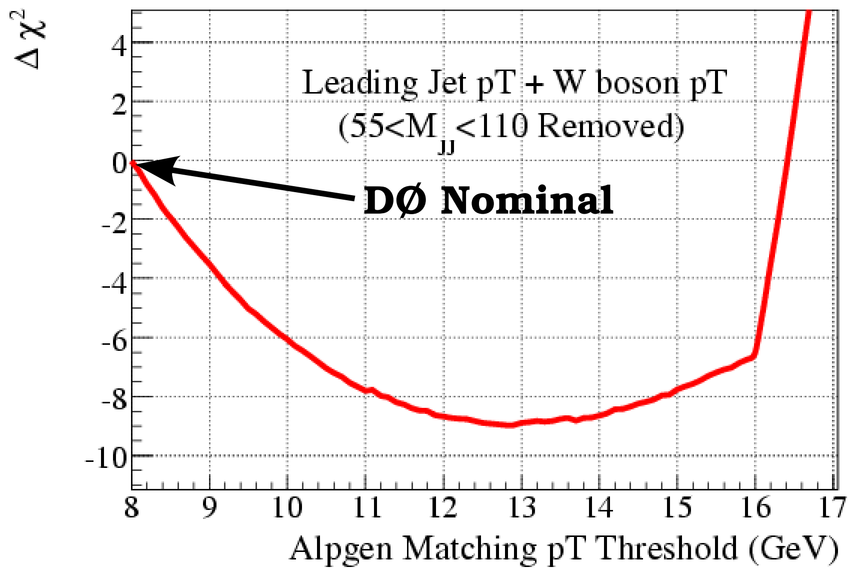
- $\Delta\chi^2$ tests show no clear preference for altered parameters, aside from MLM jet-parton matching pT threshold (*ignoring dijet mass for now*)



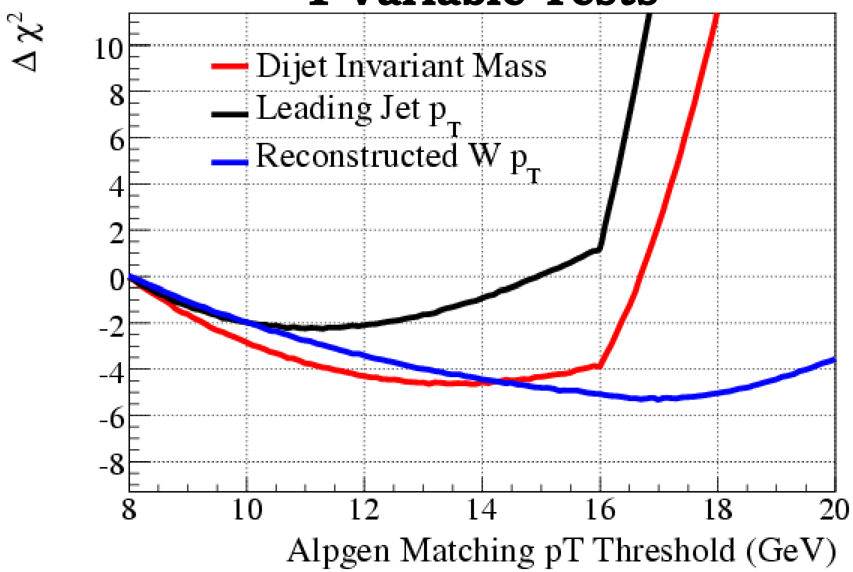
Alpgen Modeling IV

- Next, confirm these results with the full dijet invariant mass spectrum
  - Behavior is consistent in all three variables
  - Change in dijet mass at the minimum value is ~4% effect from low to high masses
  - Behavior for matching pT threshold is confirmed by Alpgen authors. Their prescription: "Generator level pT cut + 20% (or 5 GeV)" = 13 GeV for DØ.
  - *Propagate correction via event-weights*

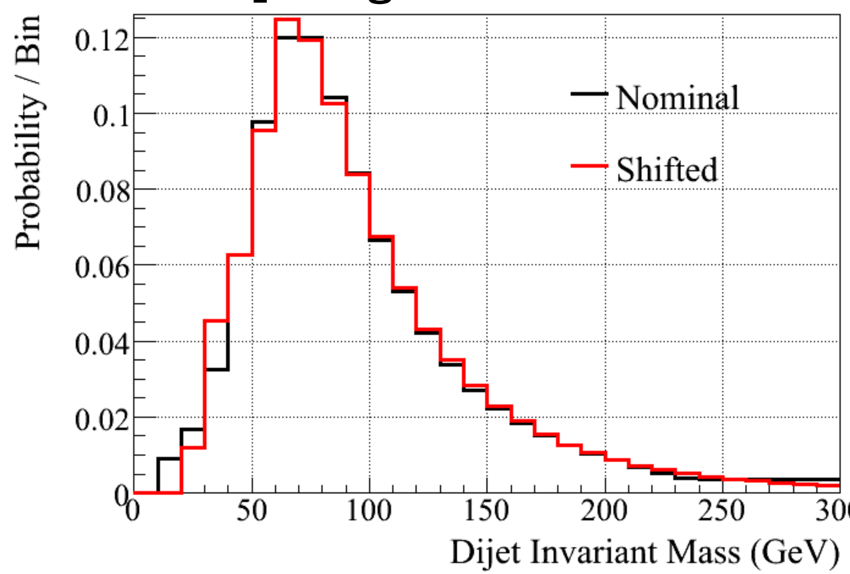
2 Variable Test



1 Variable Tests



Comparing 8 GeV with 13 GeV



Systematic Uncertainties

Source of systematic uncertainty	Diboson signal	W+jets	Z+jets	Top	Multijet	Nature
Trigger efficiency, $e\nu q\bar{q}$ channel	$+2/-3$	$+2/-3$	$+2/-3$	$+2/-3$		N
Trigger efficiency, $\mu\nu q\bar{q}$ channel	$+0/-5$	$+0/-5$	$+0/-5$	$+0/-5$		D
Lepton identification	±4	±4	±4	±4		N
Jet identification	±1	±1	±1	$\pm<1$		D
Jet energy scale	±4	±9	±9	±4		D
Jet energy resolution	±3	±4	±4	±4		N
Cross section		$\pm20^a$	±6	±10		N
Multijet normalization, $e\nu q\bar{q}$ channel					±20	N
Multijet normalization, $\mu\nu q\bar{q}$ channel					±30	N
Multijet shape, $e\nu q\bar{q}$ channel					±6	D
Multijet shape, $\mu\nu q\bar{q}$ channel					±10	D
Diboson signal NLO/LO shape	±10					D
Parton distribution function	±1	±1	±1	±1		D
ALPGEN η and ΔR corrections		±1	±1			D
Renormalization and factorization scale		±3	±3			D
ALPGEN parton-jet matching parameters		±4	±4			D

- Systematic uncertainty for all sources in units of the 1σ fractional change (%)
- The nature of each source (**N**ormalization or **D**ifferential) is also specified
- In total, 28 independent sources of systematic uncertainty

Statistical Tests

- ✗ We want to evaluate our data for the presence of a signal-like excess

  ⇒Construct a χ^2 function from the ratio of Poisson likelihoods and include prior information on the systematic uncertainties

$$\chi^2(\theta, S, B, D) = 2\sum_{i=0}^{N_{bins}} (B_i+S_i-D_i)-D_i \ln\left(\frac{B_i+S_i}{D_i}\right) + \sum_{k=0}^{N_{syst}} \theta_k^2$$

Gaussian constraint on systematics

  ⇒ can "float" systematics by removing θ^2 prior constraint (ie, free parameter in fit)

  ⇒ **For our model, both the signal and W+jets cross sections are free parameters**

- ✗ Fit Monte Carlo templates to data gives us:

  1) Best-fit signal cross section

  2) Statistical and systematic uncertainty

- ✗ Significance is obtained by fitting MC templates to pseudo-data

  Pseudo-data are drawn from background-only hypothesis (*zero signal*)

  Systematic uncertainties are randomly sampled from their prior PDFs

  Frequency of cross section outcomes at least as large as observed (or expected) cross section can be interpreted as a 1-sided Gaussian significance

Statistical Tests

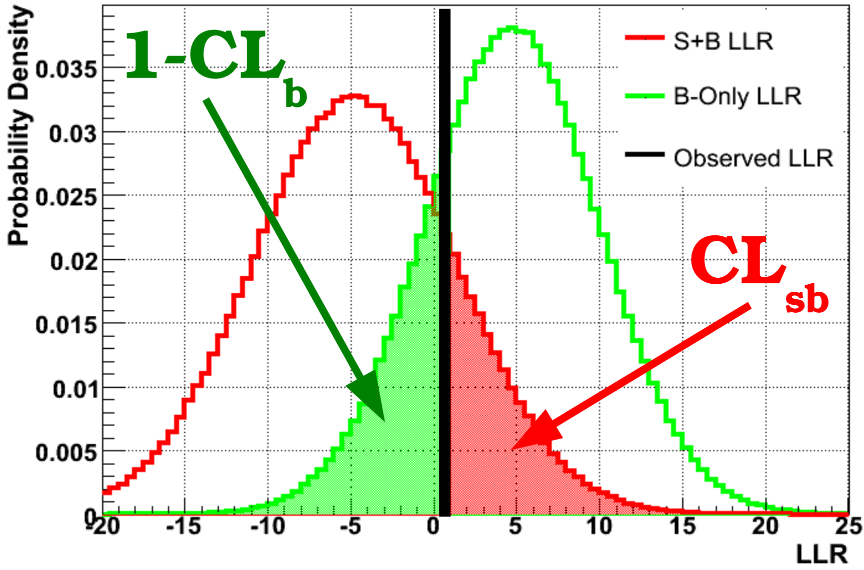
- In the case of the Higgs search, we seek to set limits on potential signal rates
  - ⇒ Similar test, comparing signal+background and background-only hypotheses
  - ⇒ Signal rate is now a fixed model parameter to be tested

$$Q = \frac{L(D|S+B)}{L^{\dagger}(D|B)} \longleftarrow$$

Two independent likelihood maximizations are performed over nuisance parameters: one for each hypothesis (S+B & B-Only)

$$LLR = -2\ln Q = \chi^2(D|S+B) - \chi^2(D|B)$$

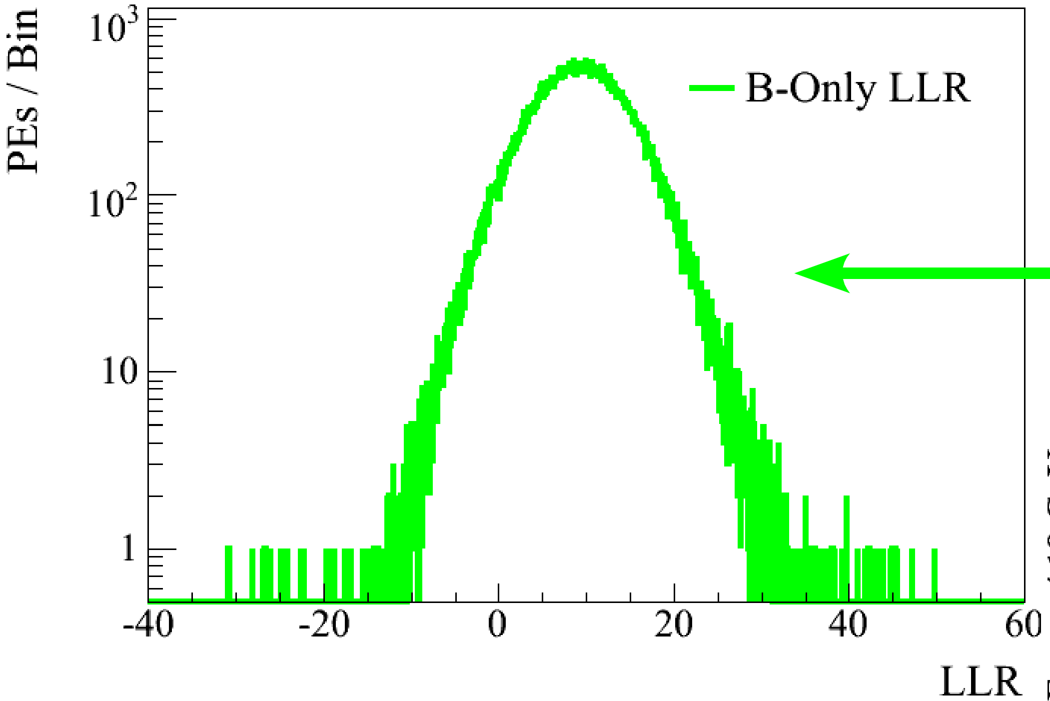
- The relative frequency of outcomes from S+B and B-Only pseudo-experiments allows us to test the signal rate
  - CLsb: fraction of S+B pseudo-experiments more background-like than data
  - CLb: fraction of B-Only pseudo-experiments more background-like than data
  - 1-CLb: fraction of B-Only pseudo-experiments more signal-like than data



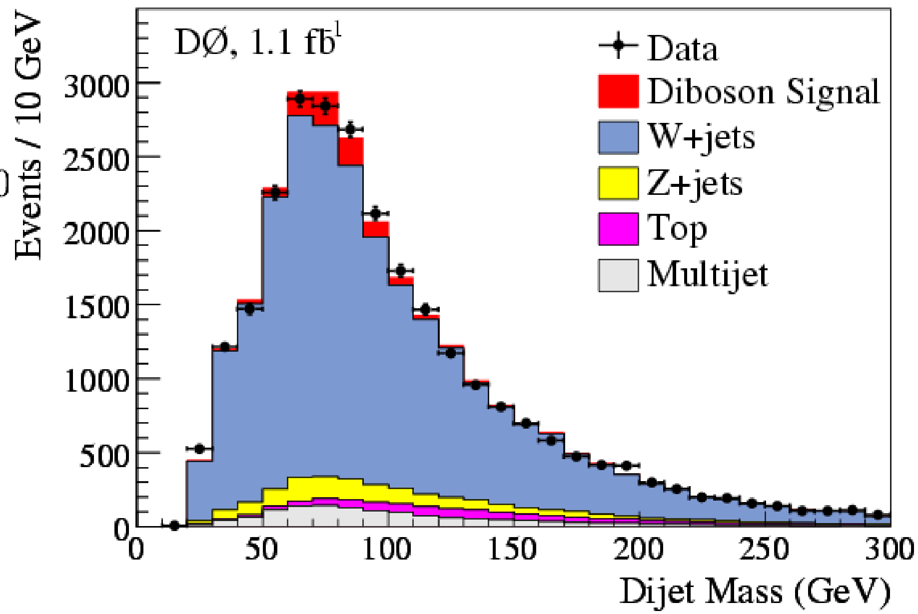
Results using Dijet Mass

✗ First, test our dijet mass templates as we would test for a Higgs signal

Are our tests consistent with a signal-like excess? Does everything work as expected?



Pseudo-Experiments drawn from the B-Only Hypothesis

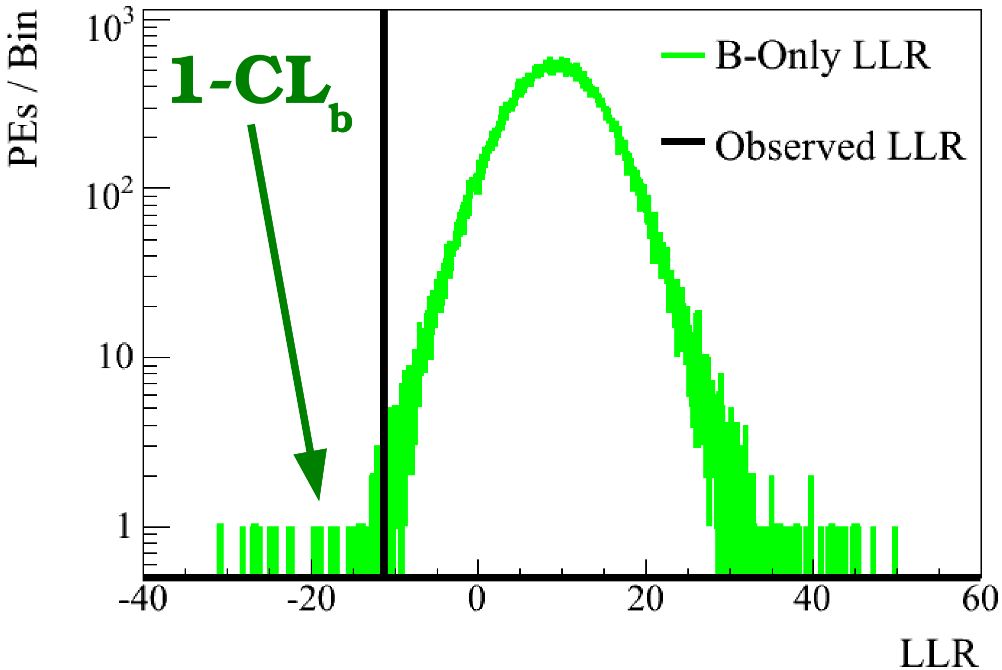


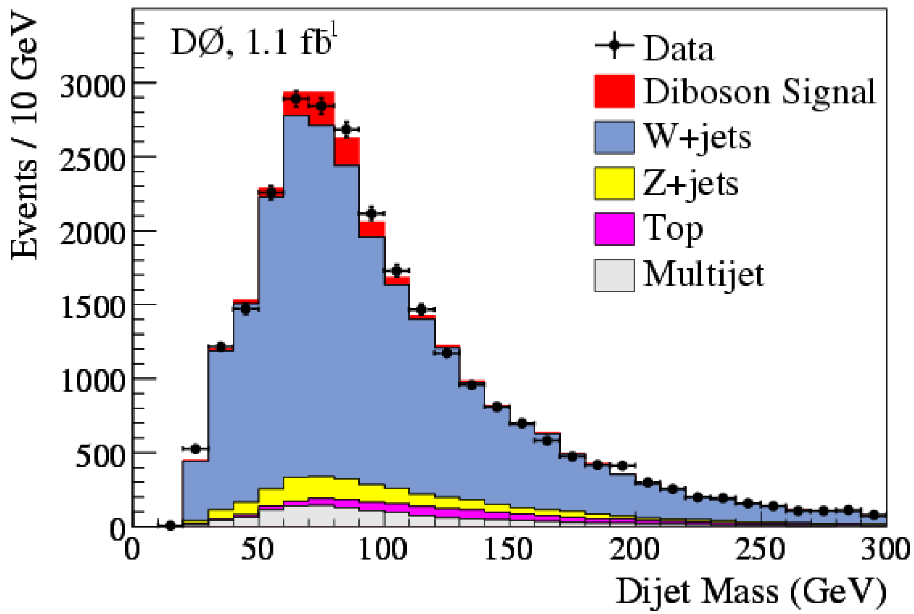
Results using Dijet Mass

- ✗ First, test our dijet mass templates as we would test for a Higgs signal

  Are our tests consistent with a signal-like excess? Does everything work as expected?

$1\text{-}CL_b = 2.5\times10^{-4}$, yielding a 1-sided Gaussian significance of **3.5 sigmas**





Results using Dijet Mass

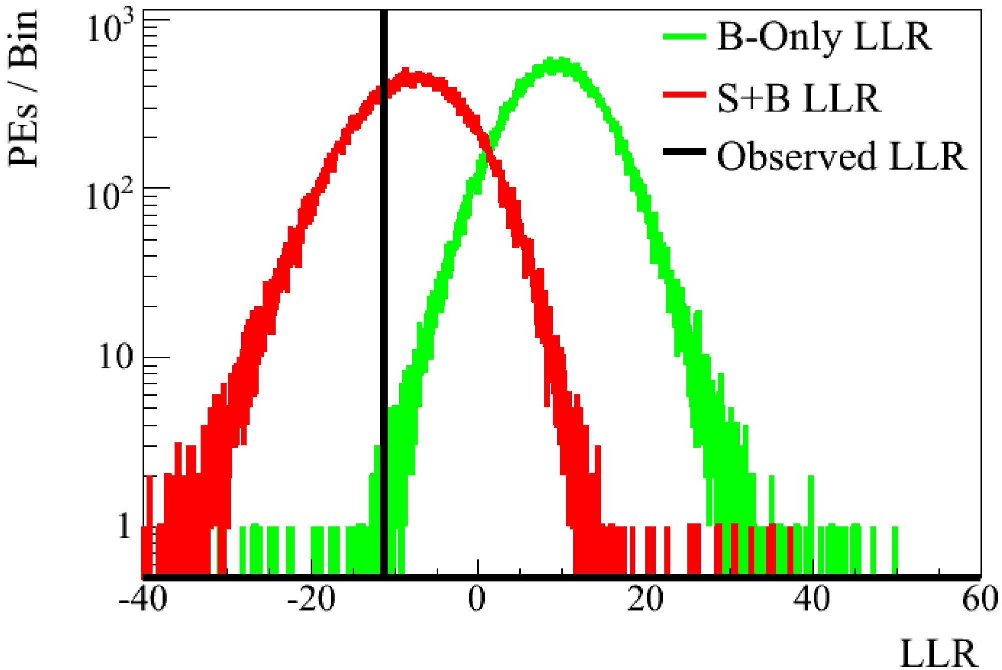
- ✗ First, test our dijet mass templates as we would test for a Higgs signal

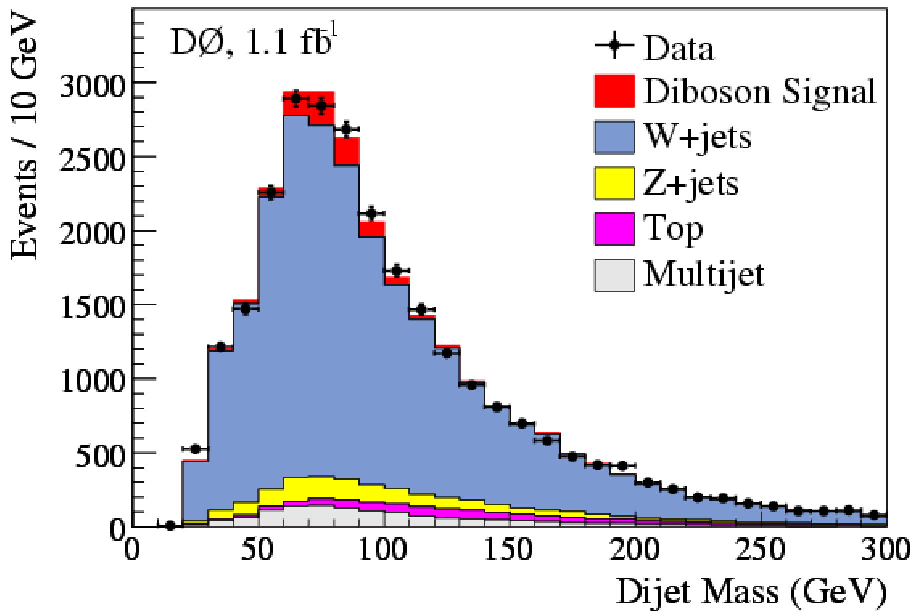
  Are our tests consistent with a signal-like excess? Does everything work as expected?

$1\text{-}CL_b = 2.5\times10^{-4}$, yielding a 1-sided Gaussian significance of **3.5 sigmas**

95% CL upper cross section limits: **9.1 pb expected, 29.7 observed**

$\sigma(WW+WZ)^{theory} = 16.1$ pb. We are sensitive to the signal and see an excess





Results using Dijet Mass II

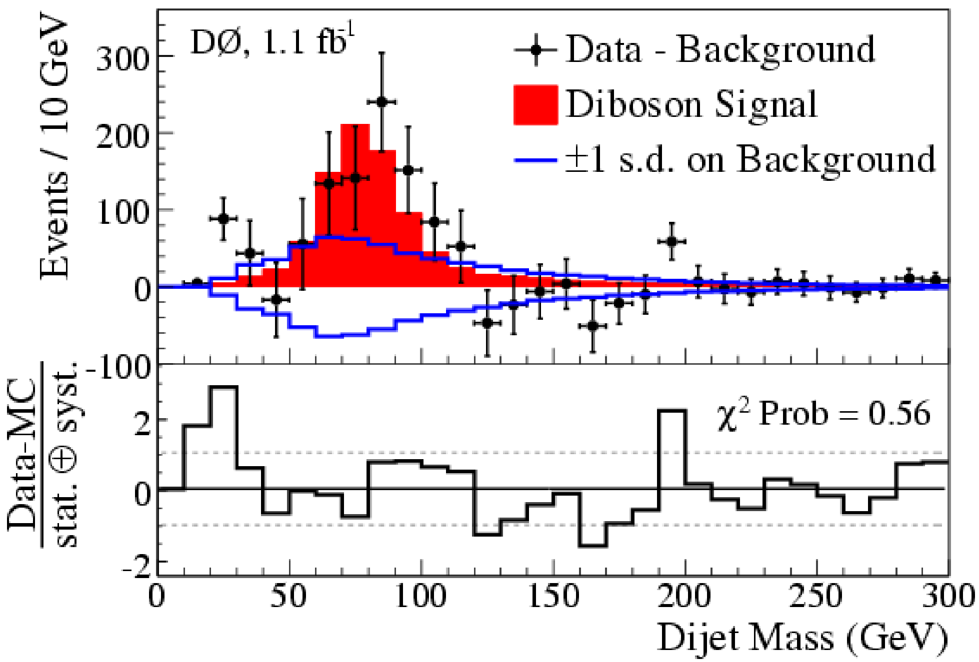
- Next, allow the signal rate to be a free parameter and fit the excess of events

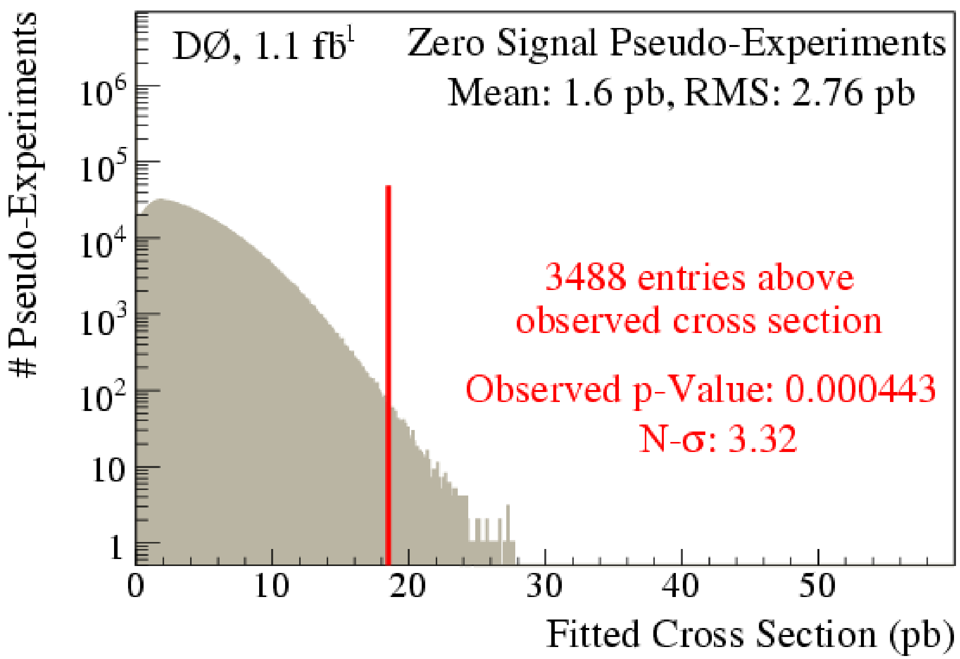
$$\sigma(WW+WZ)^{measured} = 18.5 \pm 2.8\ (stat) \pm 4.9\ (syst) \pm 1.1\ (lumi)\ pb$$

- Significance is measured by fitting the signal cross section to pseudo-data drawn from the background-only hypothesis

Expected: **2.9 standard deviations above background**

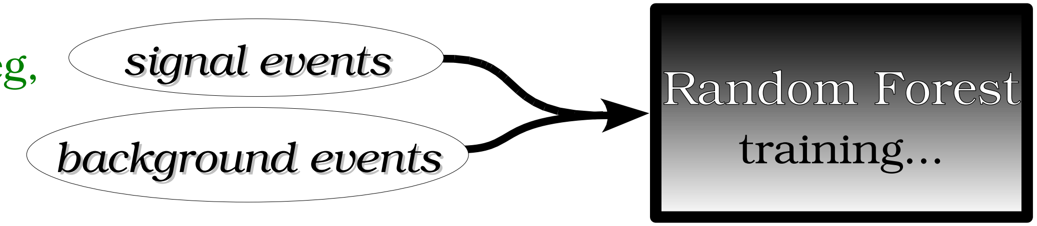
Observed: **3.3 standard deviations above background**





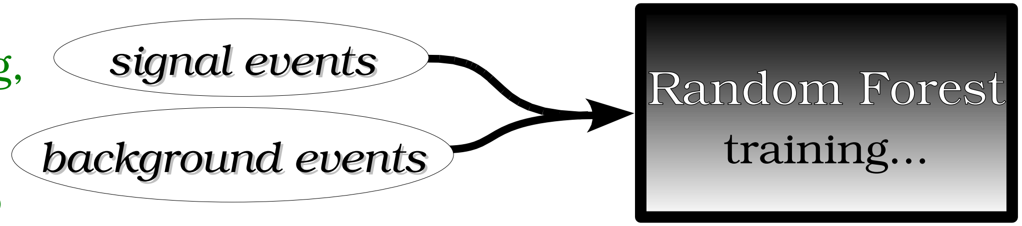
Increasing Signal Significance

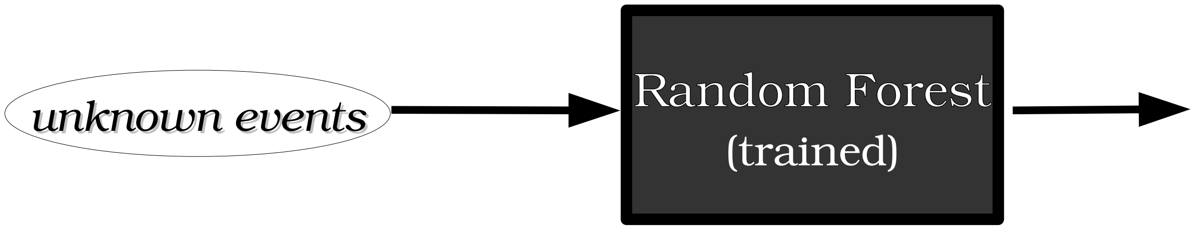
- A common technique to boost signal significance in HEP searches is to employ multivariate event classifiers. Many available options:
  - Likelihood ratios, Neural networks, Decision Trees, Matrix Elements, etc.
- We found the most powerful and robust classifier was the Random Forest
  - Our Random Forest software comes from the package "StatPatternRecognition" http://www.hep.caltech.edu/~narsky/spr.html (Ilya Narksy, Caltech)
- From the outside (**black box**), the Random Forest is similar to most machine learning techniques:
  - Train using events of known origin (eg, signal or background classes)

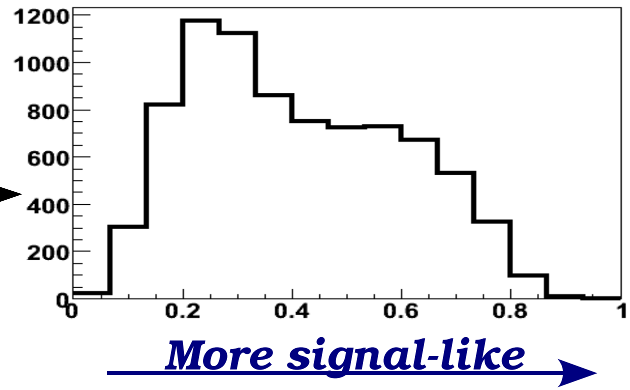


Increasing Signal Significance

- A common technique to boost signal significance in HEP searches is to employ multivariate event classifiers. Many available options:
  - Likelihood ratios, Neural networks, Decision Trees, Matrix Elements, etc.
- We found the most powerful and robust classifier was the Random Forest
  - Our Random Forest software comes from the package "StatPatternRecognition" http://www.hep.caltech.edu/~narsky/spr.html (Ilya Narksy, Caltech)
- From the outside (**black box**), the Random Forest is similar to most machine learning techniques:
  - Train using events of known origin (eg, signal or background classes)
  - The trained Random Forest is used to classify events of unknown origin as signal-like or background-like

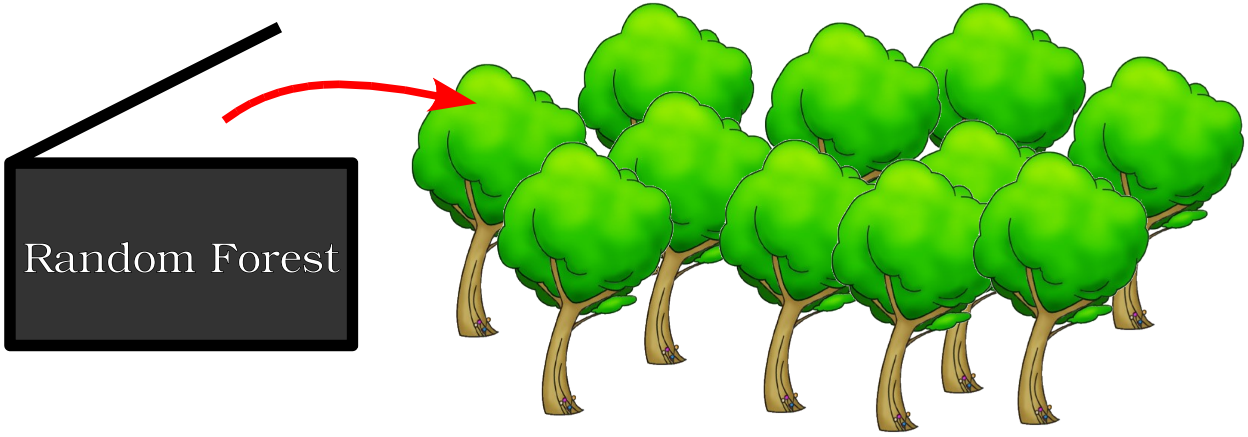


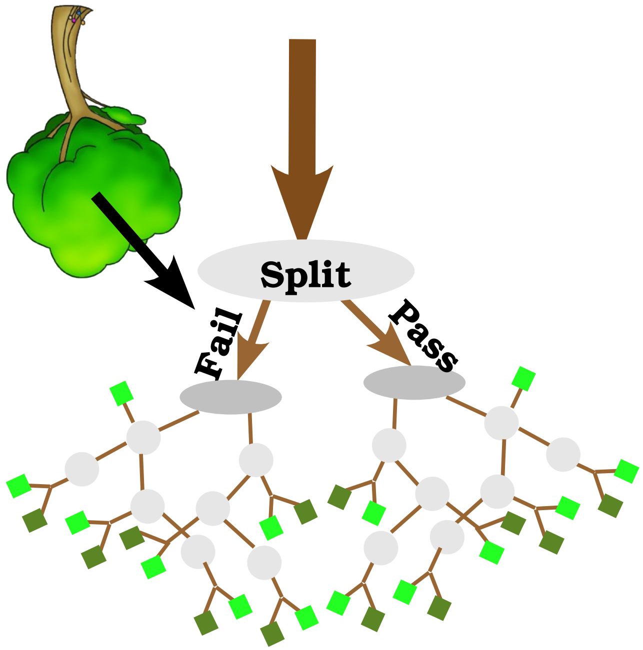




The Random Forest (RF)

A "Forest" of many decision tree classifiers

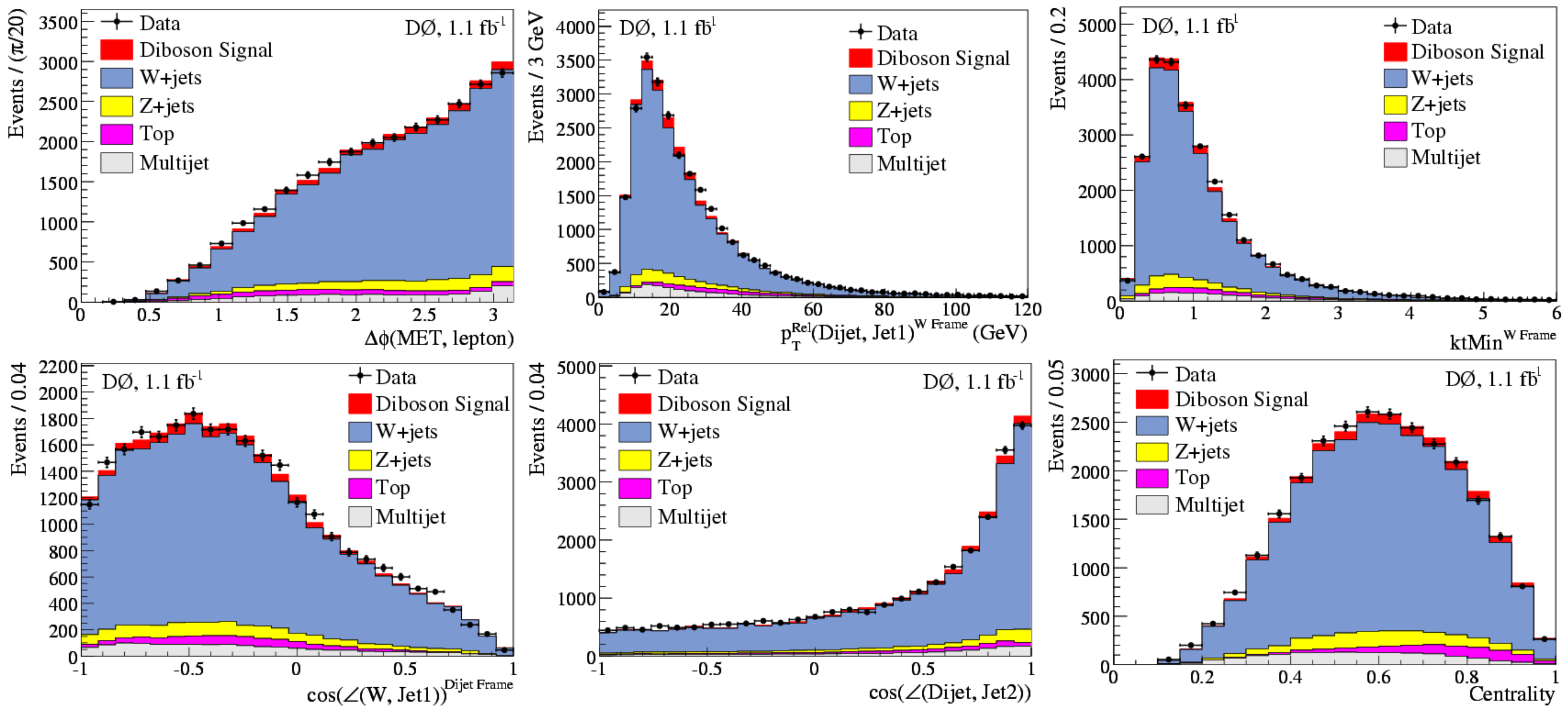




- ✗ Each decision tree in the forest is independent from all others
  - Each tree uses a random subset of the input variables
    - Allows each tree to focus on a different subset of kinematics and correlations
  - Each tree is trained using a random subset of training events
    - Provides protection against over-training and high-weight events
- ✗ The Random Forest classifier output is the **average output** over all trees
  - Fluctuations and over-training that occur for a single decision tree are reduced in the global averaging of fluctuations

RF Input Variables

- We use 13 kinematic variables as input to the Random Forest
  - Each variable helps distinguish between signal and at least one background
    - Ensure variables are well-modeled by requiring data/MC χ^2 probability outside $55 < M_{JJ} < 110$ GeV to be greater than 5%
  - In addition, include dijet mass, MET, M_T^W, jet pT, $W\rightarrow l\nu$ pT, & angular variables

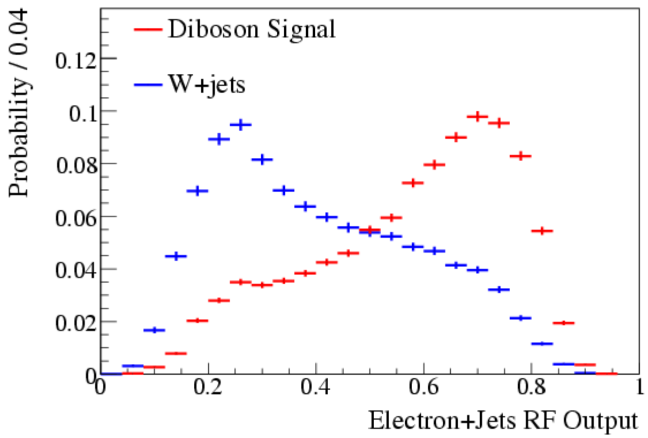


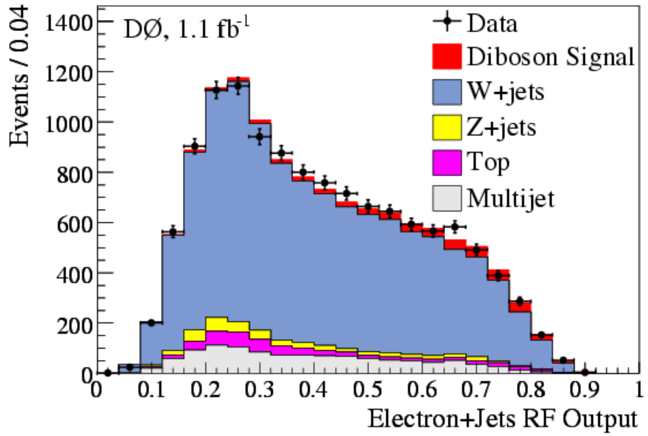
Random Forest Classifier

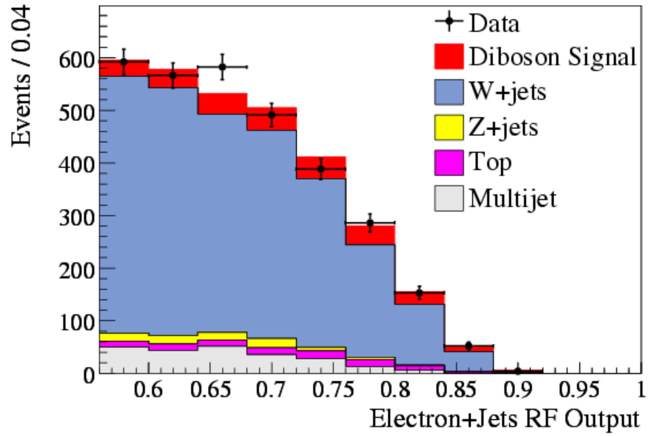
Wine & Cheese
Nov 14th, 2008

✗ The Random Forest output demonstrates improved separation of signal and backgrounds, while maintaining good agreement between MC and data

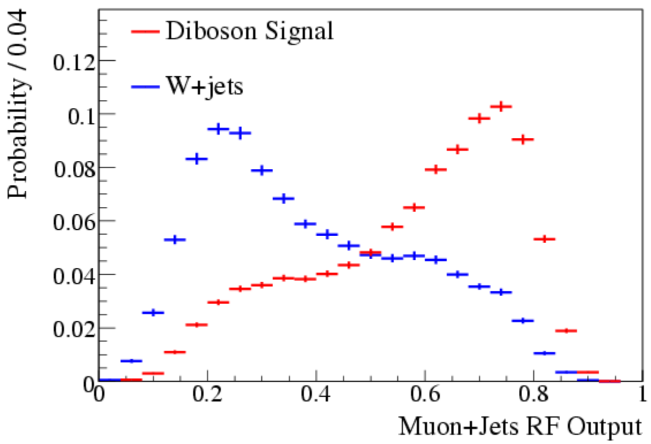
Electron + Jets Channel

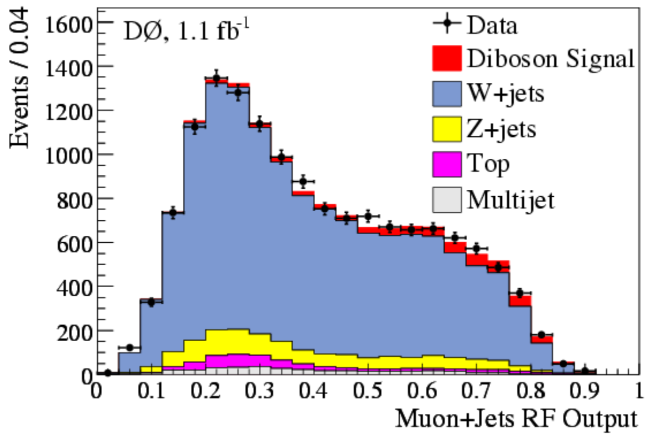


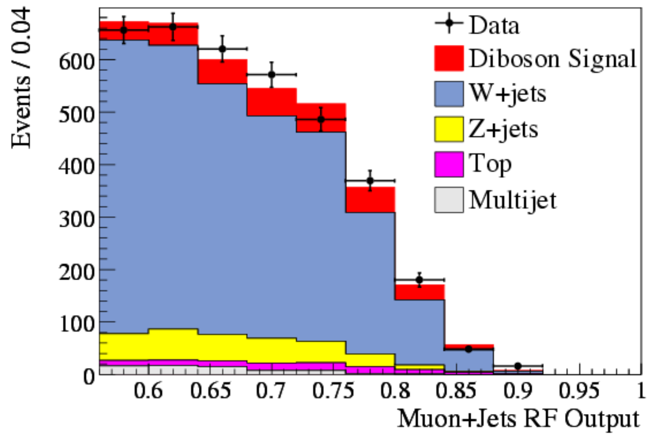




Muon + Jets Channel







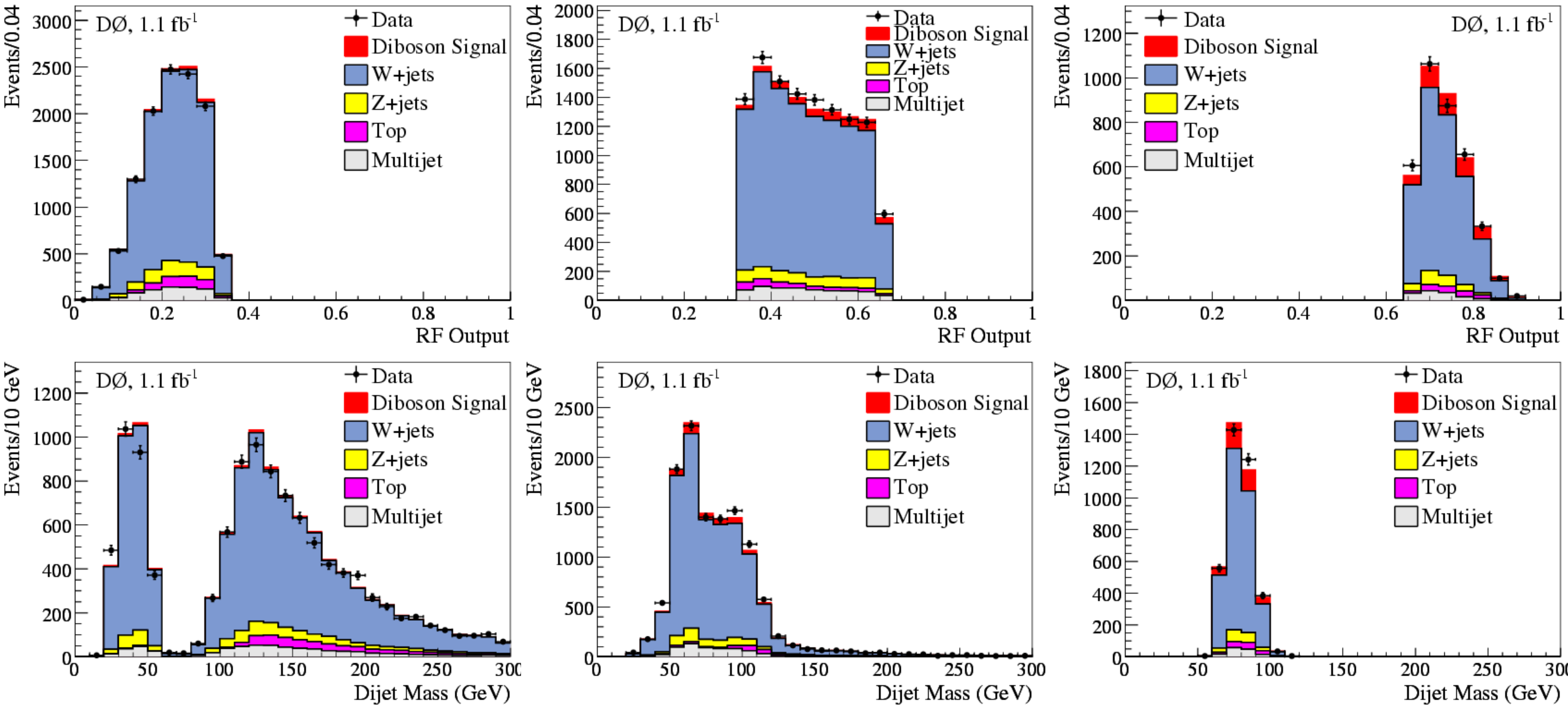
Random Forest Subspace

✗ Though the dijet mass is clearly an important variable, it is useful to understand how the RF is improving signal significance

<u>A simple test:</u> Dissect the RF to see where the events in dijet mass are going

⇒0.33<RF<0.66: the RF is efficient at moving events from the MJJ signal peak

Low Significance	**Transition Region**	**High Significance**
0 < RF < 0.33	**0.33 < RF < 0.66**	**0.66 < RF < 1**



Results with Random Forest

- Revisit our statistical tests using the Random Forest classifier

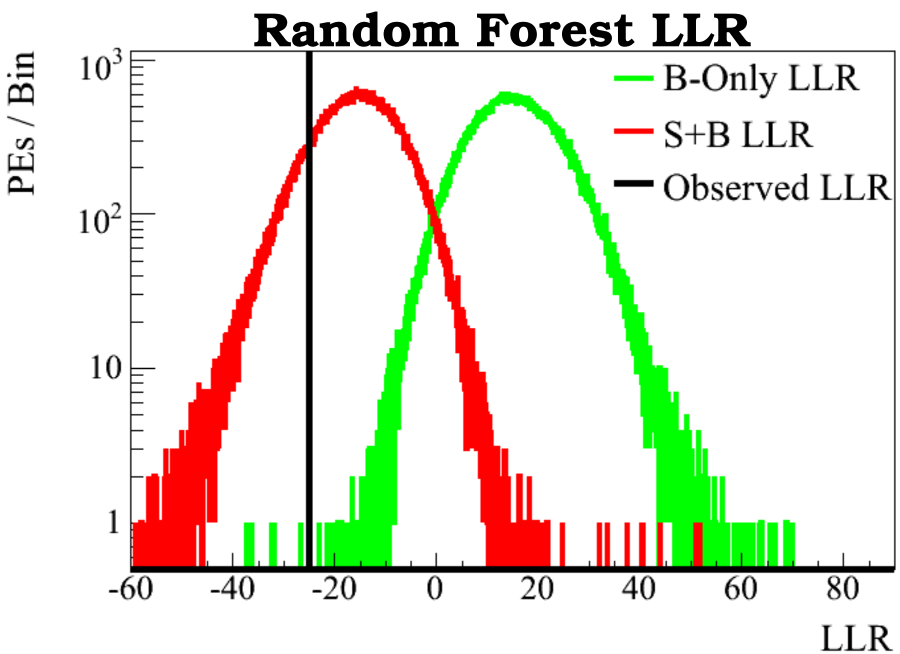
  Are we consistent with the dijet mass results? Do we improve as expected?

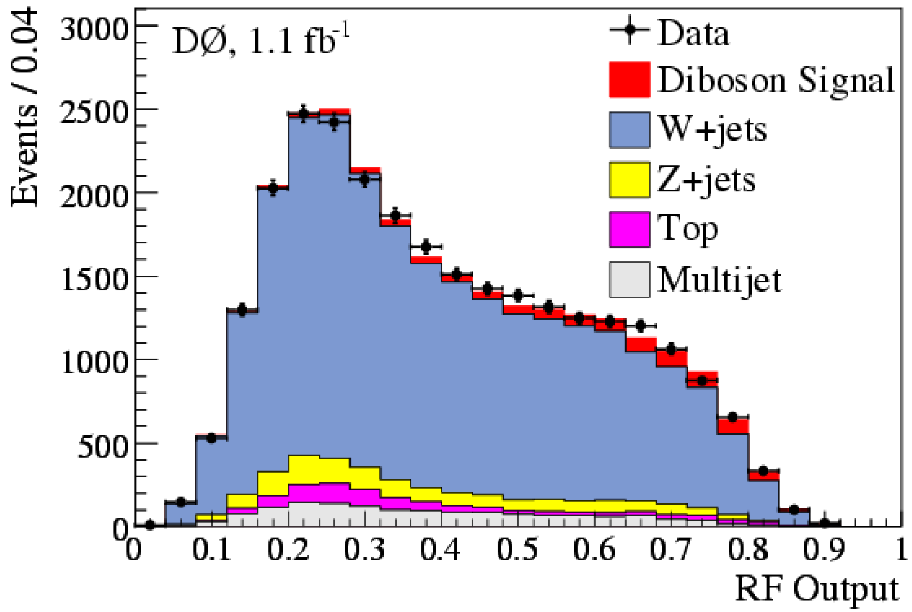
$1\text{-}CL_b = 1.8\times10^{-6}$, yielding a 1-sided Gaussian significance of **4.6 sigmas**

Dijet mass, $1\text{-}CL_b = 2.0\times10^{-4}$, or **3.5 sigmas**

95% CL upper cross section limits: **7.7 pb expected, 32.2 pb observed**

$\sigma(SM)^{theory}$ = 16.1 pb. Dijet mass: **9.1 pb expected, 25.0 pb observed**





Results with Random Forest

- ✗ Revisit our statistical tests using the Random Forest classifier

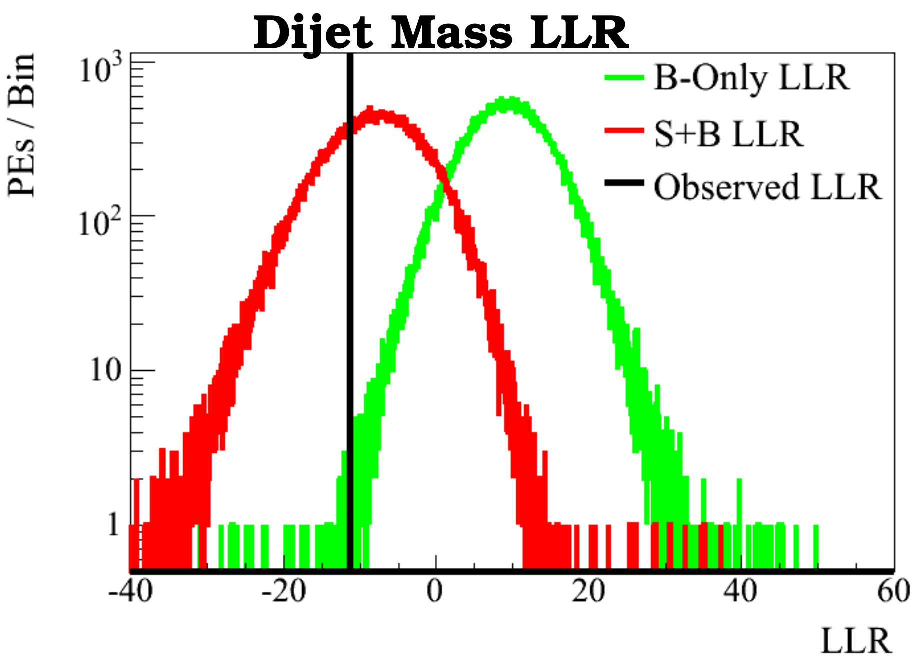
  Are we consistent with the dijet mass results? Do we improve as expected?

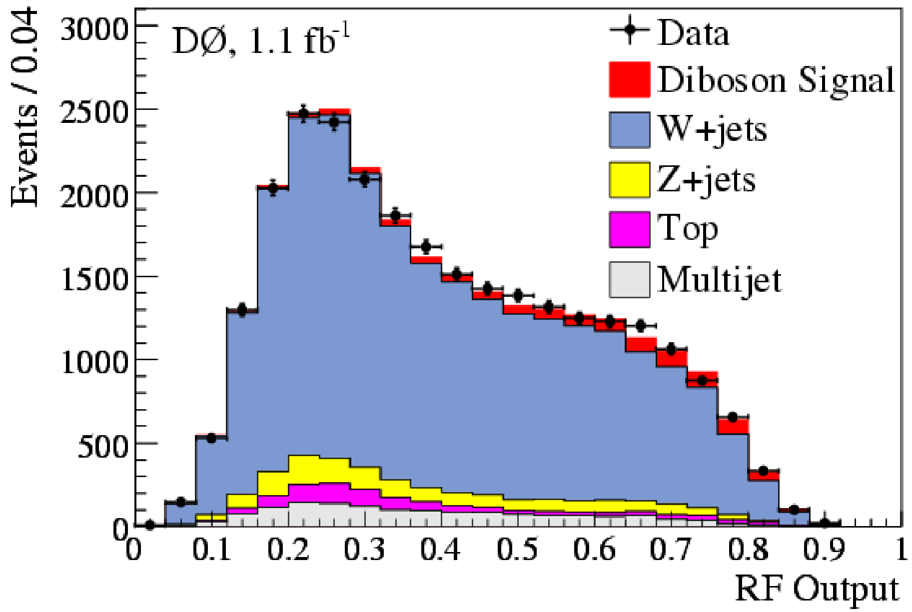
$1\text{-}CL_b = 1.8\times10^{-6}$, yielding a 1-sided Gaussian significance of **4.6 sigmas**

Dijet mass, $1\text{-}CL_b = 2.0\times10^{-4}$, or **3.5 sigmas**

95% CL upper cross section limits: **7.7 pb expected, 32.2 pb observed**

$\sigma(SM)^{theory}$ = 16.1 pb. Dijet mass: **9.1 pb expected, 25.0 pb observed**





Results with Random Forest

✗ Revisit our statistical tests using the Random Forest classifier

Are we consistent with the dijet mass results? Do we improve as expected?

1-CL_b = 1.8×10^{-6}, yielding a 1-sided Gaussian significance of **4.6 sigmas**

Dijet mass, 1-CL_b = 2.0×10^{-4}, or **3.5 sigmas**

95% CL upper cross section limits: **7.7 pb expected, 32.2 pb observed**

$\sigma(SM)^{theory}$ = 16.1 pb. Dijet mass: **9.1 pb expected, 25.0 pb observed**

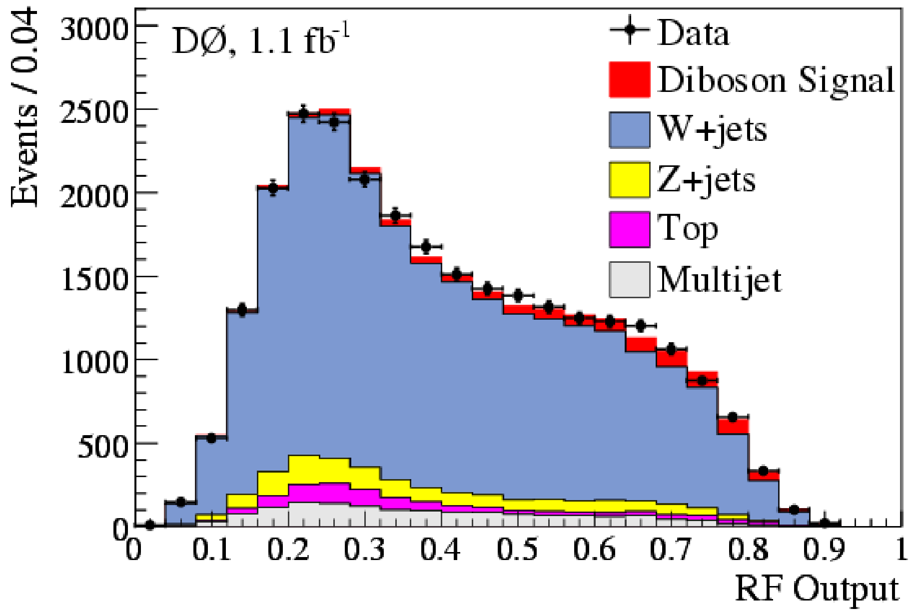
Improvements due to RF

~18% in cross section upper limit

~32% in significance

Roughly corresponds to a ~30-35% increase in effective luminosity

Expect larger Higgs search improvement



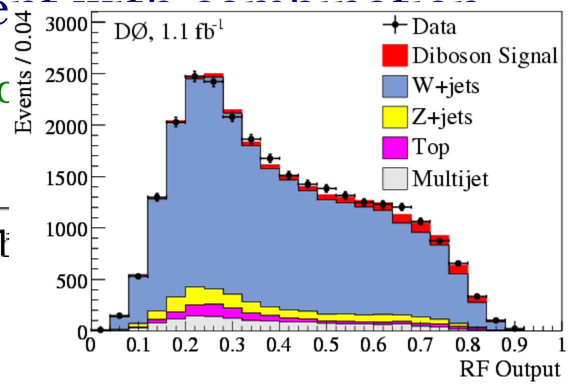
Results with Random Forest II

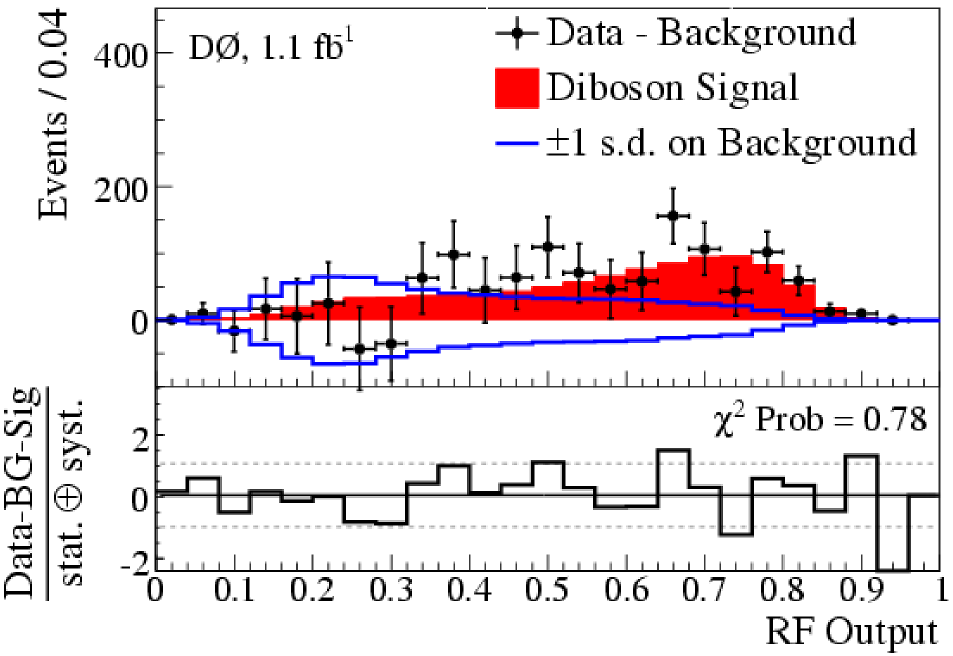
- Next, fit the Random Forest templates to data to measure the cross section

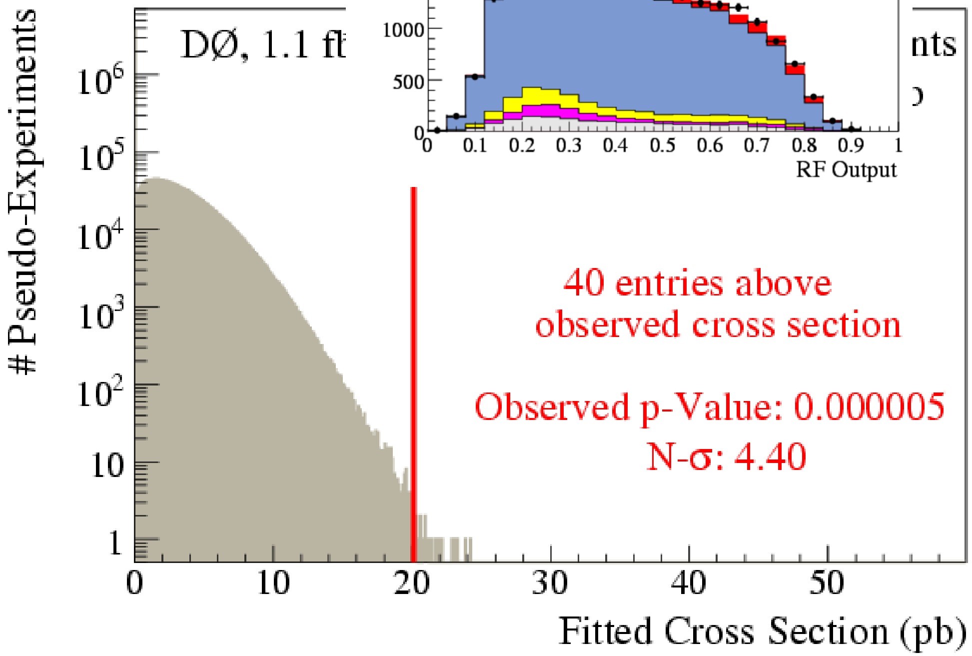
Channel	Fitted signal σ (pb)	Expected p-value (significance)	Observed p-value (significance)
$e\nu q\bar{q}$ RF Output	18.0±3.7(stat)±5.2(sys)±1.1(lum)	6.8×10^{-3} (2.5 s.d.)	3.2×10^{-3} (2.7 s.d.)
$\mu\nu q\bar{q}$ RF Output	22.8±3.3(stat)±4.9(sys)±1.4(lum)	1.8×10^{-3} (2.9 s.d.)	5.2×10^{-5} (3.9 s.d.)
Combined RF Output	20.2±2.5(stat)±3.6(sys)±1.2(lum)	1.5×10^{-4} (3.6 s.d.)	5.4×10^{-6} (4.4 s.d.)

- Electron+jets and muon+jets channels each consiste

  Larger acceptance and slightly smaller systematics in muo
  expected significance.





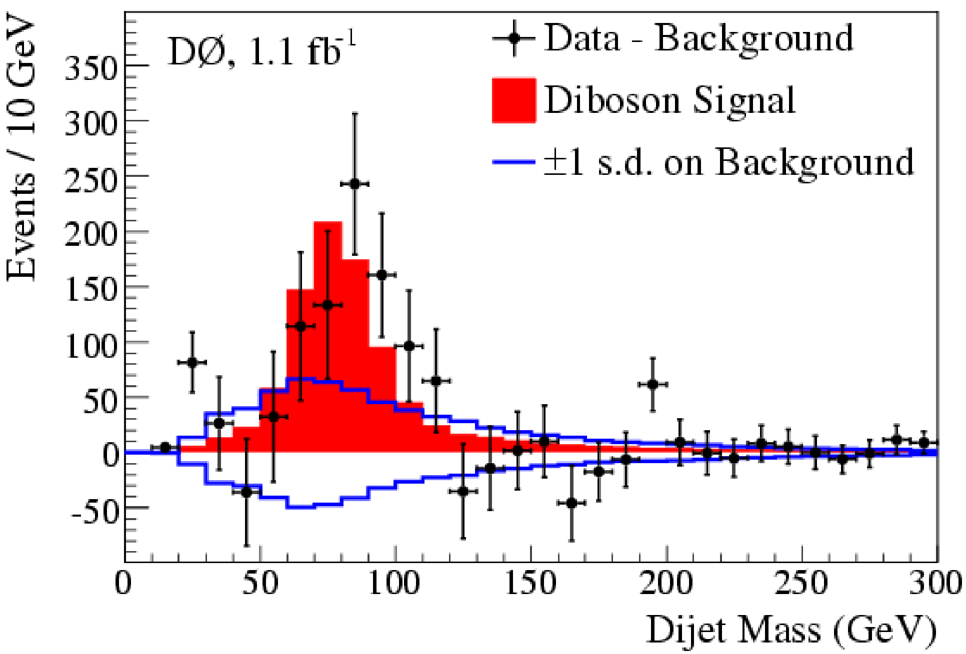


Post Fit Input Variables

✗ A final validation cross check is to check the distributions of the dijet mass after performing a fit to the RF classifier output vs a fit to the dijet mass itself

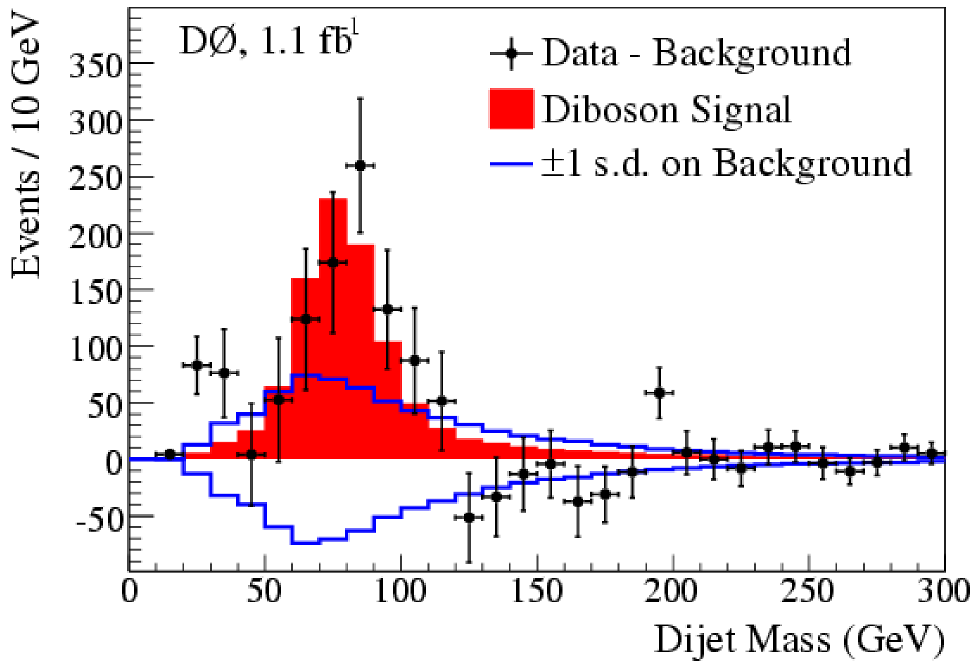
- Expect fits to achieve similar results
- Differences may indicate different biases/sensitivities in RF vs dijet mass

Dijet Mass Fit



$$\chi^2\, Probability = 0.56$$

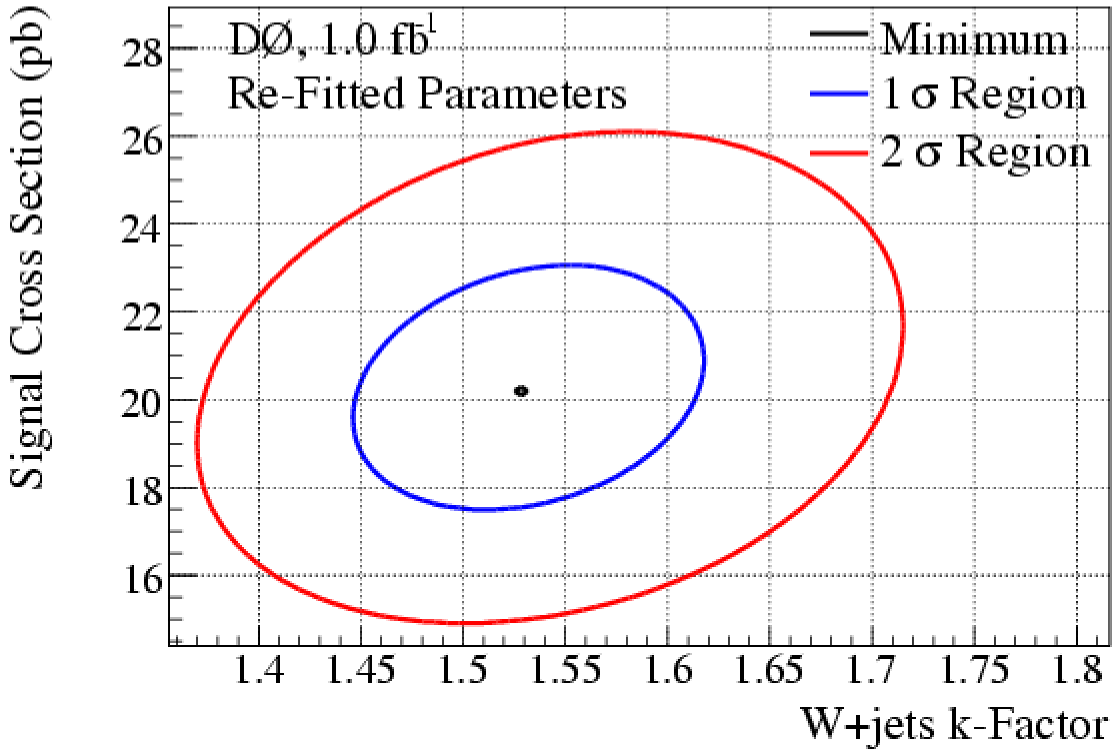
Random Forest Fit



$$\chi^2\, Probability = 0.45$$

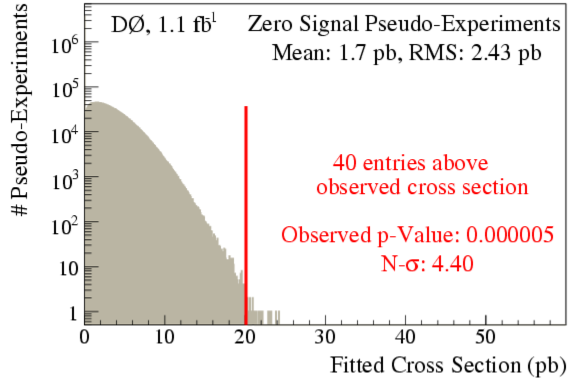
W+jets Cross Section

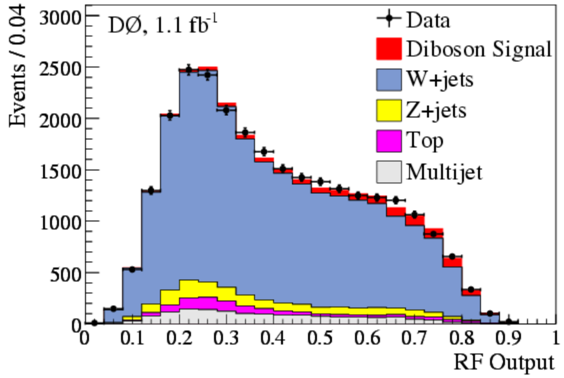
- As noted earlier, the W+jets cross section is treated as a free parameter in fits to data
  - Reduces our dependence on theoretical predictions
  - A determination of this parameter in data provides valuable feedback to other analyses using this final state (eg, WH→lνbb)
- Right: correlation ellipse for signal and W+jets cross sections
  - Remaining 27 parameters are refit for each fixed point in the plane
  - Measured value of W+jets k-Factor: **1.53±0.13**
  - Theoretical (NLO/LL) prediction: **1.52**

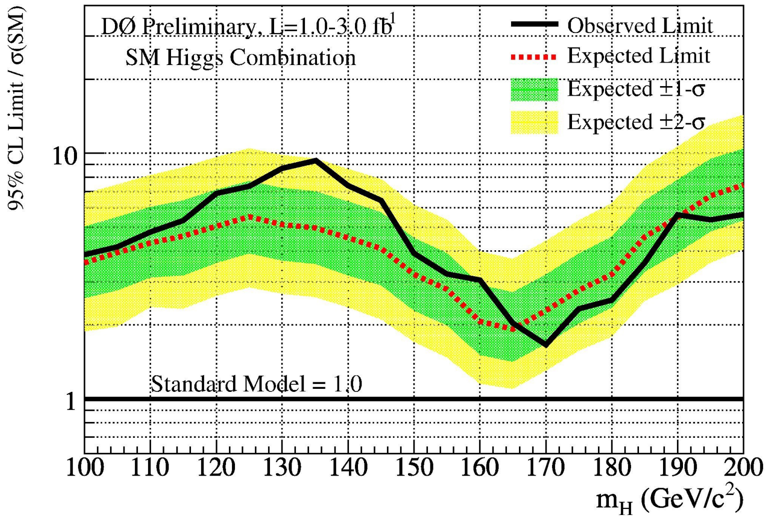


Back to the Higgs Search

- This is the first evidence of a diboson signal in a lepton+jets final state at the Tevatron
  - **Provides a crucial validation for our approach to searching for a Higgs boson**
- Statistical tools
  - We found clear evidence for a dijet mass resonance on top of a large background, keeping systematic uncertainties under control
  - Divergence of expected and observed limits led us to a **4.4 standard deviation** measurement
- Multivariate classifier
  - We demonstrated an improvement in the significance of a real signal by introducing the Random Forest classifier
  - An effective increase in luminosity of ~35%
- The Tevatron Higgs search
  - These tests don't guarantee we can find the Higgs, but they provide a **proof of principle**







Conclusions

- We present the first evidence for WW/WZ→lνjj decays at the Tevatron

$$\sigma(WW+WZ)^{measured} = 20.2 \pm 2.5\,(stat) \pm 3.6\,(syst) \pm 1.2\,(lumi)\,pb$$

$$\sigma(WW+WZ)^{theory} = (\underbrace{12.4}_{\textbf{WW}} + \underbrace{3.7}_{\textbf{WZ}}) \pm 0.1\,pb = 16.1 \pm 0.1\,pb$$

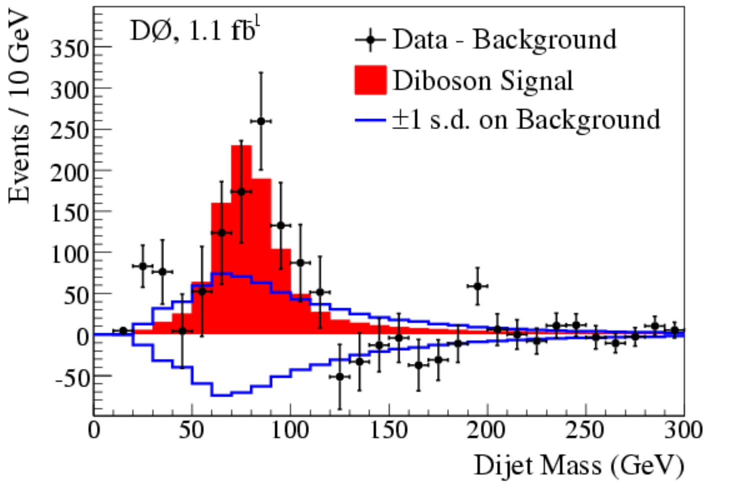
Signal significance: **4.4 standard deviations** above background

Submitted to PRL

Available at: Fermilab-Pub-08/457-E, arXiv:/0810.3873 [hep-ex]

- Results are consistent with previous Tevatron diboson measurements in purely leptonic final states

- DZero's new data sample has reached 4 fb^{-1}
  - Higher precision cross section measurement
  - Separation of WW and WZ final states
  - Anomalous triple gauge coupling limits



Backup Slides

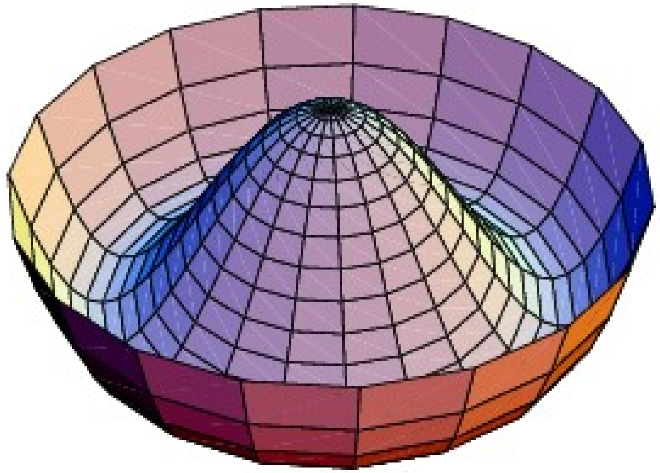
Connecting with Higgs Theory

- Just by talking about weak vector boson masses, we're already talking about the Higgs

$$M_W^2 = \frac{1}{4} g^2 v^2 \qquad M_Z^2 = \frac{1}{4}(g^2 + k^2) v^2$$

Remaining particle masses are put in "by hand"

$$V_{Higgs} = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2$$

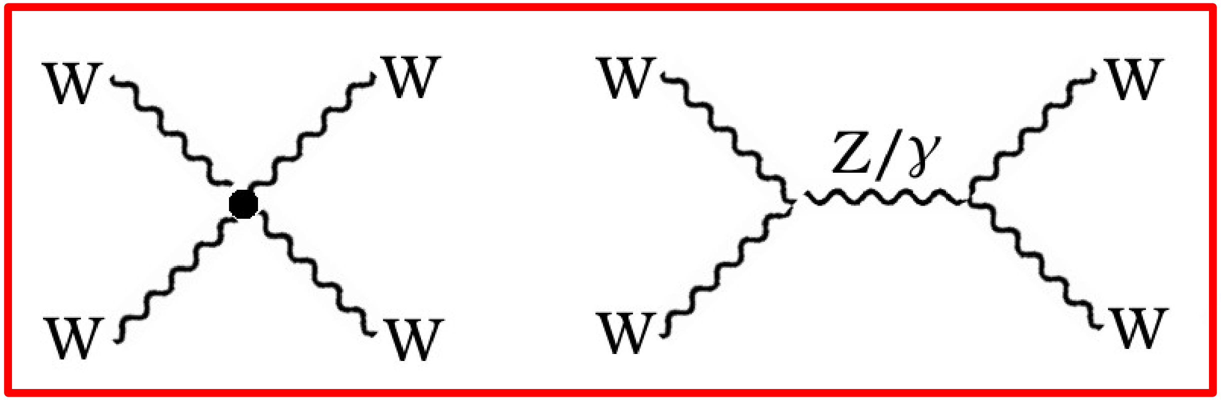


$$\phi_0 = \pm\sqrt{-\mu^2/2\lambda} = v/\sqrt{2}$$

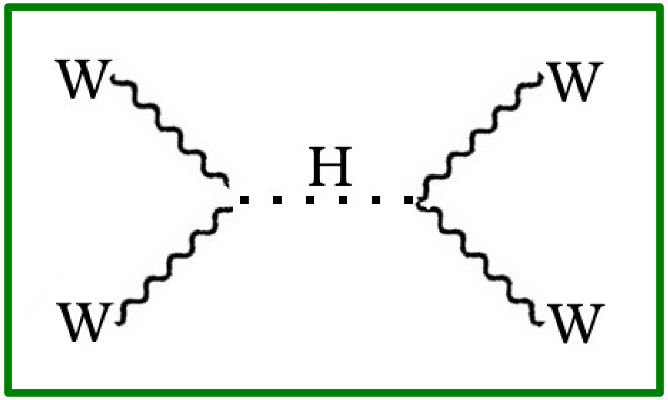
- Furthermore, the Higgs comes to the rescue again to maintain unitarity in WW scattering

True for $M_H < \sim 900$ GeV

Cross section diverges like s/M_W^2



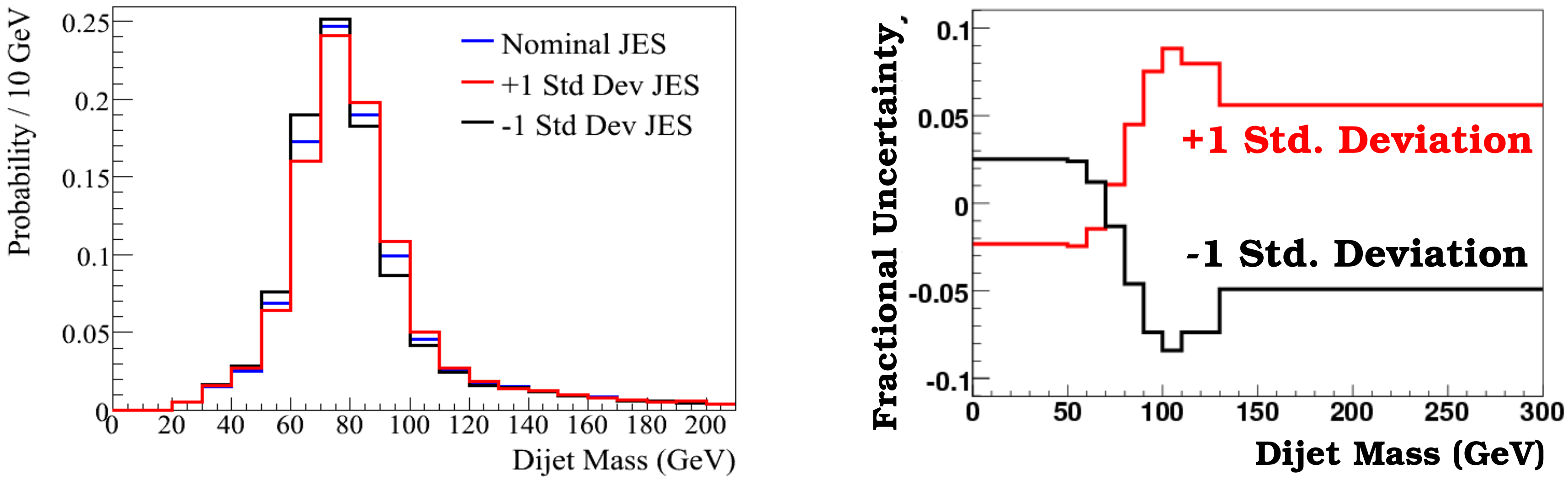
Scalar Higgs cancels divergence



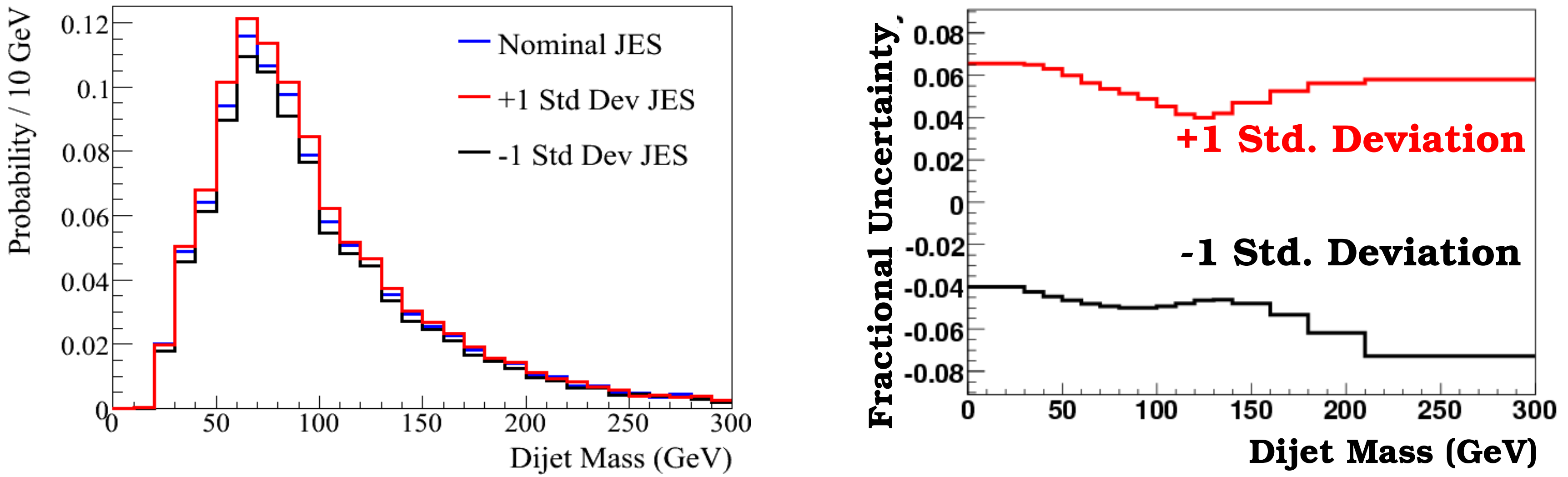
Systematic Uncertainties

- ✗ One of the dominant sources of systematic uncertainty is the jet energy scale

Jet energy scale uncertainty for signal events



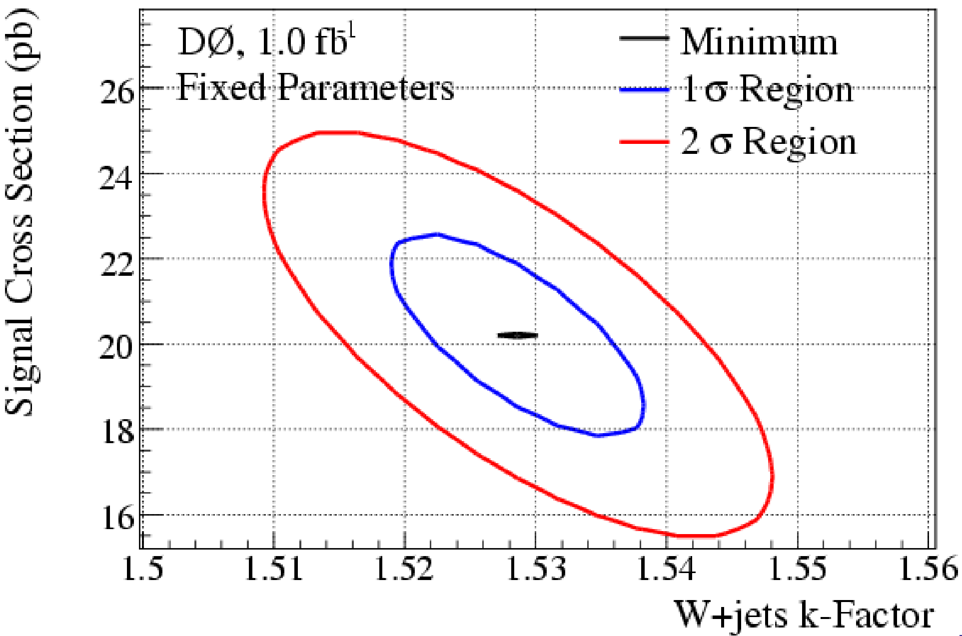
Jet energy scale uncertainty for W+jets events



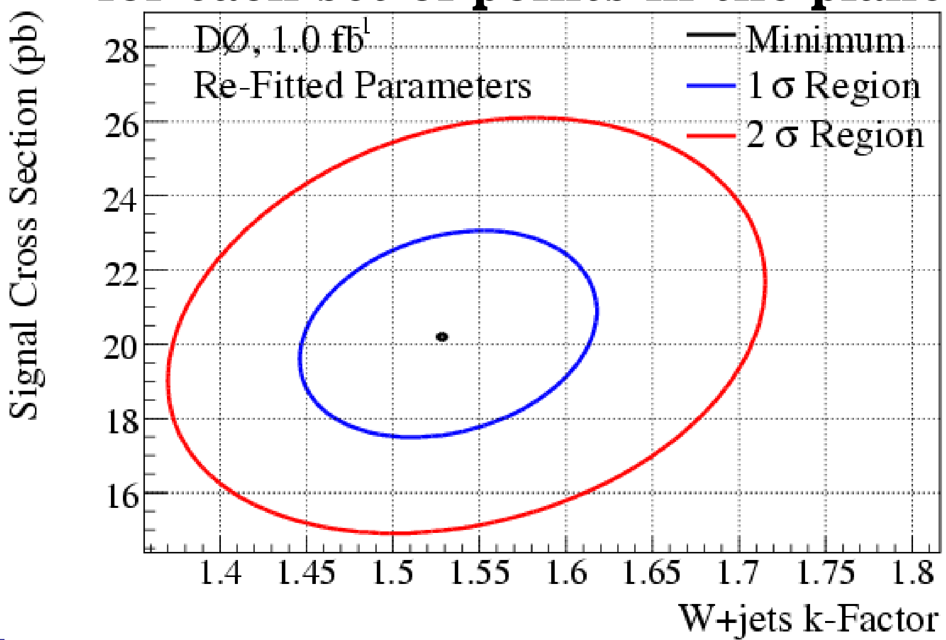
W+jets Cross Section

- The W+jets cross section is treated as a free parameter in fits to data
  - Reduces our dependence on theoretical predictions
  - Correlations can be extracted in two ways:
    1) Fix all parameters at values from minimum. Demonstrates compensating correlation effect of cross sections
    2) Refit all parameters for each set of points. Reflects actual measurement and reveals size of remaining degrees of freedom

Remaining 28 parameters fixed at values from minimum

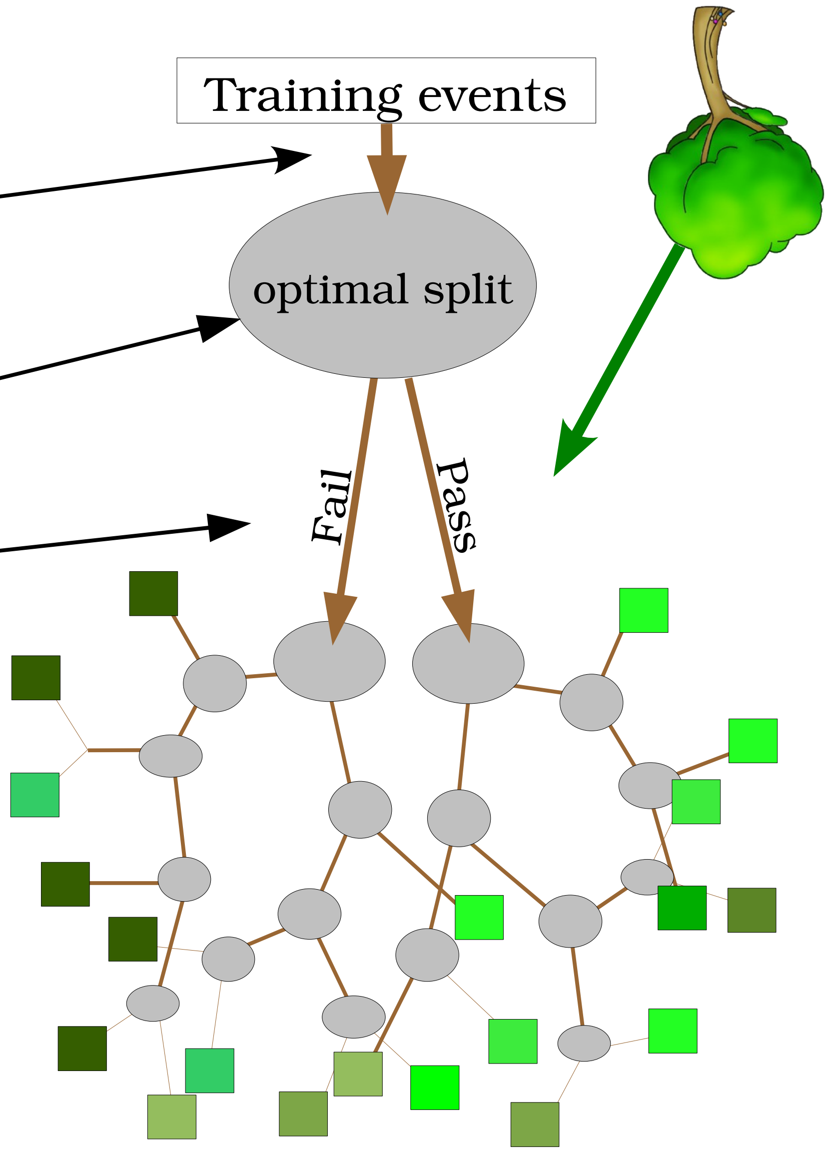


Remaining 28 parameters re-fit for each set of points in the plane



Decision Trees

- Decision tree is trained/grown using a set of known signal & background training events
  - ⇒ These events go into the root node
- Algorithm looks at all possible splits on all input variables and applies split giving best separation between signal and background
- Events pass into one of two child nodes depending on whether they pass or fail
- This process is repeated until:
  - A node contains all or no signal events
  - # events per node is less than a pre-specified amount (optimized for each application)
- Output for an unknown event is determined by the signal purity of the terminal node that the event ends up in



Statistical Tests

- ✗ With our carefully studied Monte Carlo and data samples, we may now proceed to perform statistical tests of the system.

  Aim to validate Frequentist search techniques with a "real" signal

  Our basic goal: *Use data to distinguish two competing hypotheses:*

  H0 ⇨ null hypothesis: background-only model (zero signal)

  H1 ⇨ a test hypothesis: background + signal

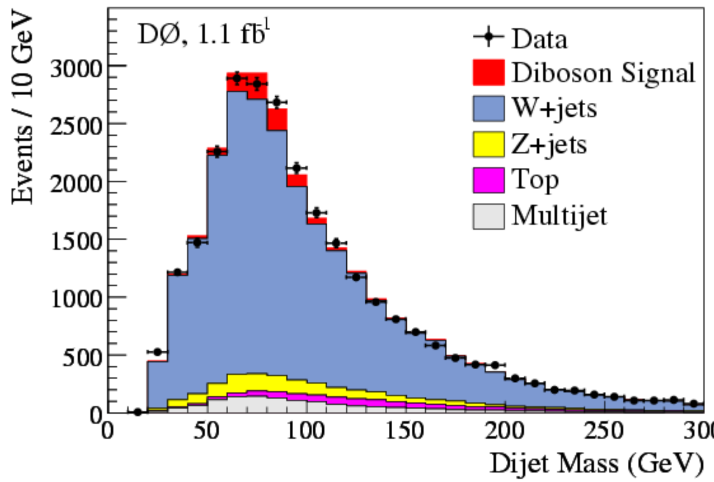
  H0 is a compound hypothesis with some set of nuisance parameters

  H1 has the same form, but add extra nuisance and model parameters

- ✗ Following the Neyman-Pearson lemma, we construct a test statistic (ordering rule) based on the relative joint likelihoods for the hypotheses:

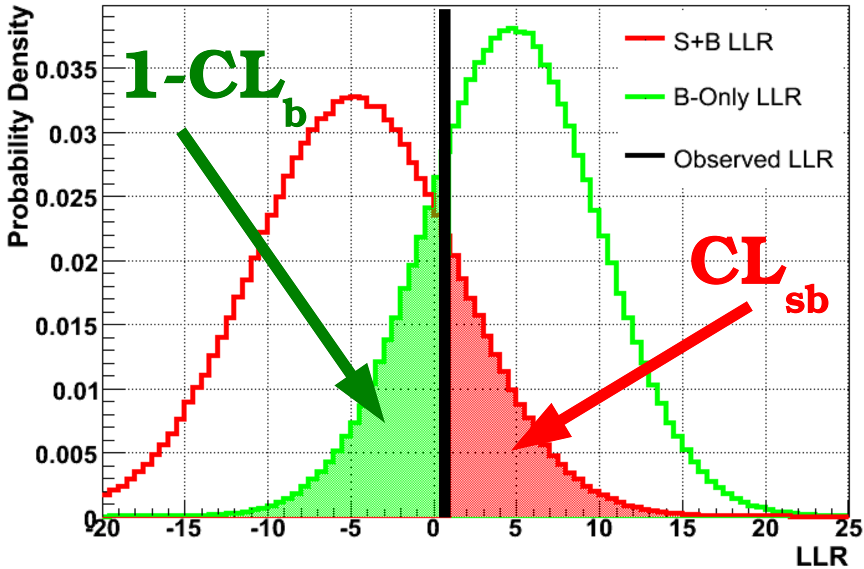
$$Q = \prod_{j=0}^{N_{bins}} \frac{L_{ij}(H1)}{L_{ij}(H0)}$$

We can treat each bin of a histogrammed distribution as a semi-independent test



Statistical Tests II

- ✗ The Frequentist approach:
  - Assume data is drawn randomly from a Poisson parent distribution
    - ⇒ Generate pseudo-data via random Poisson trials with mean value taken from expected backgrounds (**H0**) or signal-plus-background (**H1**)
- ✗ A complication: Our hypotheses are compound (*they contain nuisance parameters*) and we need to test a simple hypothesis
  - Systematics are a tricky Frequentist problem, so use a Bayesian model
    - ⇒ Model uncertainties on nuisance parameters as Gaussian-distributed, sample randomly for each pseudo-experiment (*including correlations*)
- ✗ Distribution of test statistic for each PE defines "prior predictive ensembles"
- ✗ **CL_b** (**CL_{sb}**)= fraction of **HO** (**H1**) outcomes less signal-like than data
- ✗ Width of distributions arises from Poisson ⊕ systematic uncertainties



Statistical Tests III

- Our system (*also the Higgs search*) is systematics-limited: signal≈3% of background, uncertainty≈20% of background.
- Solution: Counteract the degrading effects from uncertainties via "Profile Likelihood" technique

  Likelihood now a function of signal, bkgd, data, and nuisance parameters

  Maximizing the likelihoods for a set of data points defines our "best fit" for that data (or pseudo-data) in a given hypothesis

$$Q = \frac{L(x|\theta_{R1}, \hat{\theta}_S)}{L(x|\theta_{R0}, \hat{\hat{\theta}}_S)}$$

Two independent likelihood maximizations are performed over nuisance parameters parameters, one for each pseudo-experiment

θ_{R1} θ_{R0} : Physics parameters in ***H1*** and ***H0***, respectively

$\hat{\theta}_S$ $\hat{\hat{\theta}}_S$: Nuisance parameters which maximize L for ***H1*** and ***H0***, respectively

Result: data-constrained systematics, narrowed Q distributions, *improved separation of **H1** and **H0***

Statistical Tests IV

- ✗ Profile Likelihood: defined by a fit of our MC model to data

Assume prediction of N events per bin is a function of nuisance parameters

$$\hat{B}_i \rightarrow B_i \prod_{k=0}^{N_{syst}} (1+\sigma_i^k S_k)$$

B_i = nominally predicted bin content
σ_i^k = fractional uncertainty
S_k = N sigma deviation from nominal

Assume data is Poisson distributed, derive joint likelihood over histogram bins and include prior information on systematics

$$\hat{B}_i \rightarrow B_i \prod_{k=0}^{N_{syst}} (1+\sigma_i^k S_k)$$

$P_i(X)$ = Poisson PDF for X events in bin i
G(Y) = Gaussian PDF for systematic K

By considering -2lnQ, the resulting X2 is linear in bins

Ratio of Gaussian priors reduces to $S^2 = \left(\frac{\hat{\sigma}-\sigma^0}{\sigma^0}\right)^2$

⇒ can "float" nuisance parameters by removing S^2 prior constraint

$$\chi^2 = -2\ln Q = 2\sum_{i=0}^{N_{bins}} (\hat{B}_i - D_i) - D_i \ln\left(\frac{\hat{B}_i}{D_i}\right) + \sum_{k=0}^{N_{syst}} S_k^2$$